Evaluation of targeted methylation at the CDKN2B promoter

Lamadema, Nermina

Awarding institution:
King's College London

Evaluation of targeted methylation at the *CDKN2B* promoter

Nermina Lamadema

University of London

A THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

KING'S COLLEGE LONDON

2012

Declaration

I hereby declare that I alone composed this thesis and that the work presented here is my own, except where stated otherwise.

March 2012

Acknowledgments

I would like to begin by thanking the Leukaemia Research Fund for supporting me with the funding during my PhD. Without their financial support my project would not have been possible.

To my supervisor Dr Kevin Ford whose patience, encouragement and kind support led me to the threshold of understanding the subject, thank you for the teaching, for giving me the opportunity to learn from you and for your encouragements during these “targeting” years.

I have also been indebted to my second supervisor whose academic experience, sound advice, and inspirational ideas have been invaluable to me.

Special thanks to Professor Farzin Farzaneh and Professor Ghulam Mufti for additional funding when I needed it most.

It has been a pleasure knowing and working with my lab colleagues and friends past and present, whose comradery I’ll forever cherish. I offer my sincerest thanks to those who supported me in any respect during the completion of the project.

Finally, to my lovely family, my husband Vaslav, my daughter Nikki, my mother and brother words can’t express the gratitude I feel for all your support, and all the sacrifices you made during these years. You are my pillars of strength, and my guiding light and it is to you that I dedicate this thesis.

Table of Contents

Abstract

DNA 5- Cytosine methylation is a modification that occurs in mammalian cells predominantly at CpG sequences, through addition of a methyl group to the C5 position of cytosine. It is a commonly held view in the current field of epigenetics that the aberrant methylation patterns observed in gene promoter regions are associated with transcriptional repression in the absence of genetic mutations. This epigenetic phenomenon is frequently reported in malignant haematopoiesis, where genes involved in cell growth and differentiation programmes are shown to be transcriptionally silenced leading to unchecked cellular proliferation.

Although strongly implicated in gene shut down, it is still unclear whether promoter region methylation triggers gene repression, or if DNA methylation arises as a consequence of transcriptional gene silencing *via* some other mechanism. This is in part because no-one has yet managed to dissect *in vivo* the precise sequence of the events associated with the appearance of aberrant methylation patterns at the promoter region of actively transcribing genes.

This thesis documents a research programme aiming to exogenously impose *de novo* cytosine methylation to specific regions of a relevant endogenous promoter in order to study spatial and temporal interactions between targeted DNA methylation and other components of the epigenetic regulatory network. The gene selected for this study was *CDKN2B*, a cell cycle regulating tumour suppressor gene strongly implicated in Myelodysplastic Syndrome - MDS and Acute Myeloid Leukaemia - AML. Site-biased zinc finger: DNA methyltransferase fusion proteins were specifically designed with the aim of depositing low and high density *de novo* methylation at different regions of the *CDKN2B* promoter, in an attempt to simulate methylation patterns associated with disease and to thus determine their effect on gene expression.

We demonstrate here for the first time the establishment of low density *de novo* methylation at this endogenous locus, characterized by highly specific target site methylation downstream of zinc finger binding sites. Examination of the distribution pattern and the spread of methylation from the seeded methylation hotspots suggest that 5' CpG positions in flanking sequences remain unmethylated, whilst 3'CpG acquire *de novo* methylation marks derived from the action

of endogenous DNA methyltransferases. The effects of methylation spread over longer time scales on the formation of broader and more uniform methylation patterns are currently being investigated. Our data also shows that once established, the methylation pattern is inherited through successive cell divisions. Gene expression analysis indicates sustained transcriptional activity of the low density methylated promoter despite the presence of these *de novo* methyl marks. Overall, our results demonstrate that targeted low density methylation to the promoter region of the *CDKN2B* gene can be established and maintained epigenetically through subsequent cell generations, but has no effect on the transcriptional regulation of the gene. To date we have not been successful in targeting high density methylation to the same region to comprehensively evaluate the bilateral relationship between a densely methylated promoter and its transcriptional activity/regulation.

The contribution of the transcription factor VEZF1 in mediating methylation protection at the *CDKN2B* promoter was also examined, and we showed for the first time that this transcription factor is strongly associated with this promoter *in vivo*. This may have important implications in our understanding of *CDKN2B* expression and regulation in the disease process. Related studies also revealed that TET proteins may play a significant role in cell cycle entry, which was another hitherto unknown phenomenon.

In conclusion, we feel that the work in this thesis has advanced the field of targeted methylation, the further implementation of which will certainly further our understanding of key epigenetic mechanisms such as the cellular response to *de novo* methylation and its role in the regulation of an important cell cycle and disease associated gene.

Chapter 1

General Introduction

1.1 Mechanisms of Epigenetic Regulation

Epigenetics is defined as “The study of mitotically and meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence [1].” DNA methylation is one well characterized mechanism for propagating epigenetic information, although other mechanisms do exist such as chromatin remodelling events and miRNA silencing machinery. It is unclear which of these mechanisms are primary and which secondary epigenetic marks. The relationship between chromatin remodelling complexes and the DNA methylation machinery appears to be a dynamic one. DNA demethylation is typically associated with activated chromatin, whilst DNA methylation is more reflective of a constituently silenced chromatin state [2]. Chromatin architecture in terms of nucleosome space occupancy around gene transcriptional units also contributes significantly to the complexity of epigenetic gene regulation.

Research has demonstrated that both environmental and nutritional factors influence the epigenome markedly *via* mechanistic components of the DNA methylation and chromatin modification machinery. The processes of gene expression, extracellular signalling and cellular differentiation are linked, regulated and defined by the epigenetic mechanisms of chromatin signature/DNA methylation interplay. Recent findings suggest that an organisms methylome may be the way by which it responds to, and adapts to natural environmental fluctuations [3].

Epigenetic changes on their own or combined with genetic alterations are believed to directly initiate cancer development [4]. A comparison of the key epigenetic differences between cancer and normal genomes shows that the biggest changes occur in the DNA methylation pattern between the two states. Similarly, alterations in patterns of covalent histone modifications can deregulate the control of chromatin-based processes again, resulting in oncogenic transformation. The above epigenetic processes are potentially reversible which given the important role they play in cancer, makes them viable targets for therapeutic intervention.

1.1.1 DNA Methylation

DNA methylation is one of the earliest discovered epigenetic modifications. In 1975, Riggs, [5] and Holliday and Pugh [6] proposed a role for DNA methylation in cellular memory as a mechanism for X-chromosome stable silencing early on in development. Analysis of the establishment and maintenance of differential gene expression patterns in many other developmental processes has further implicated DNA methylation in transcriptional gene silencing in a broader context. These processes include early embryonic *Hox* gene expression [7], differential expression of maternal and paternal copies of the genes (genomic imprinting) [8], and haematopoiesis. Roles for DNA methylation in the regulation and maintenance of many cellular processes, chromosome stability, and maintenance of chromatin structure and transcription control have also been established.

Gene and tissue - specific methylation patterns established early on in embryogenesis ensure stable propagation of the epigenetic code. X -chromosome inactivation by *de novo* methylation mentioned above, is an example of this mechanism [9]. Transcriptional reactivation of genes involved in cellular differentiation programmes occurs through active demethylation processes [10].

A key feature of the mammalian methylome is global hypermethylation, observed most frequently in highly conserved, heterochromatic, repetitive regions such as satellite DNA, short and long interspersed transposable elements and endogenous retroviral sequences. Methylation occurs in 70-80% of CpG sites in this region and serves as a mechanism for keeping the region transcriptionally inert [11].

Exceptions to this state of global hypermethylation are CpG rich islands (CGIs) - short regulatory regions, usually 1kb in size, organized in asymmetric clusters, usually located in the 5' end of the gene, and within the promoter [12]. CGIs are methylation free in all normal tissue types, irrespective of gene transcription status. There are a few exceptions such as the regulatory regions of the genes involved in X inactivation [13], spermatogenesis [14] and the imprinted genes *H19* and *Igf2r* [15, 16]

Changes to DNA methylation patterns occur throughout the life of an individual. These changes can arise as a consequence of nutritional effects [17], and age - related cellular degeneration [18, 19]. Aberrant methylation patterns implicated in the aetiology of cancer frequently associates with global hypomethylation and promoter region hypermethylation. Disruption of DNA methylation patterns, in highly repetitive regions such as transposon repeats or viral genes can lead to global genomic instability [20], resulting in neoplastic transformation. This may also arise as a consequence of disruptions in transcriptional control of tumour suppressor genes, oncogenes, and genes linked to DNA repair machinery.

These changes in DNA methylation status correlate with the inhibition of gene expression, with a potential knock on effect on cellular differentiation and proliferation, cell cycle control, genomic imprinting and X chromosome inactivation [21]. However, evidence against a strong association between DNA methylation and gene silencing also exists, where for example several genes have shown a positive co-relation between DNA methylation and tissues specific expression [22-25].

Like most epigenetic modifications, DNA methylation is a reversible process, subject to cyclic changes in gene promoter regions [26]. Global methylation changes during fertilization and early development further support the idea that DNA methylation is a dynamic process, while active demethylation which mediates global methylation loss from the parental genome in mammalian fertilized oocytes before the onset of DNA replication reinforces this view [27, 28]. In preimplantation embryos, mammalian blasto-cysts acquire *de novo* methylation patterns [29]. The immune response can also be regulated by methylation dynamics. For example, studies on the mouse *Interleukin-2* gene promoter enhancer region during T-cell activation, shows that the methylation state of a specific, single CG site is a key regulatory mechanism [30].

Thus, the global distribution of DNA methylation patterns in human somatic cells and its dynamic regulation through the interplay with a variety of epigenetic factors ensures a fluid epigenomic state that responds to stimuli by continuous remodelling events.

1.1.2 Mechanisms of DNA methylation

The C5 carbon of cytosine bases within the palindromic 5-'CpG-3' dinucleotides are the predominant targets of covalent methyl group addition. The primary DNA sequence does not change as a result of the chemical modification.

DNA methylation is mediated by a family of DNA methyltransferase enzymes (DNMTs) using S-Adenosyl methionine (SAM or AdoMet) as the universal methyl group donor.

The C6 position is the site of covalent DNA methyltransferase association during CH_3 group transfer, whilst S-adenosylhomocysteine (SAH) is the by- product of the reaction and is inhibitory at high concentration. Methionine adenosyltransferase catalyses the generation of SAM in an enzymatic reaction involving L-methionine and ATP [31]. Ultimate transfer of the CH3 group occurs from the SAM methionine sulphur atom to cytosine base at C5 of CpG in DNA. After methyl group transfer, SAH dissociates from the complex (Figure 1-1).

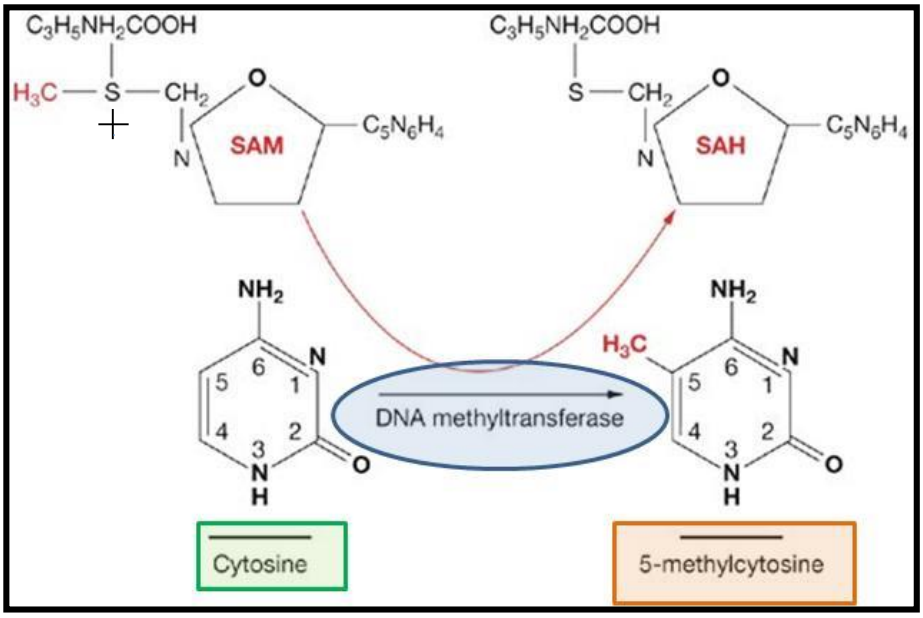


Figure1-1. Schematic diagram of DNA methylation process.

DNA methylation involves transfer of a methyl group (CH_3) from SAM onto the C5 position of cytosine residue. The process is catalysed by a group of enzymes called DNA methyltransferases DNMT which attacks the C6 position of the cytosine residue during the CH_3 transfer. Figure adapted from *Nature -Clinical Practice Rheumatology* (2007), B.Richardson.

Eukaryotic DNA methylation is catalysed by the DNA methyltransferase (DNMTs) family of enzymes. Initial methylation patterns are established by *de novo* DNMTs which act on unmethylated DNA. The family consists of three catalytically active members: the DNMT1 maintenance methyltransferase, and DNMT3A/ DNMT3B involved in *de novo* methylation. DNMT3L is catalytically inactive. Prokaryotic DNA methyltransferases are smaller than their more evolutionary advanced eukaryotic relatives. A major difference is in N-terminal domain of eukaryotic DNMTs, which contain a zinc ion binding, and cysteine rich region [32], called the plant homeodomain (PHD) or ATRX-like domain with similarity to the PHD region of *ATRX* gene, which is an essential ATP- dependant chromatin associated protein (Figure 2). This observation has led to suggestions that the DNMT family also plays a role in chromatin remodelling events.

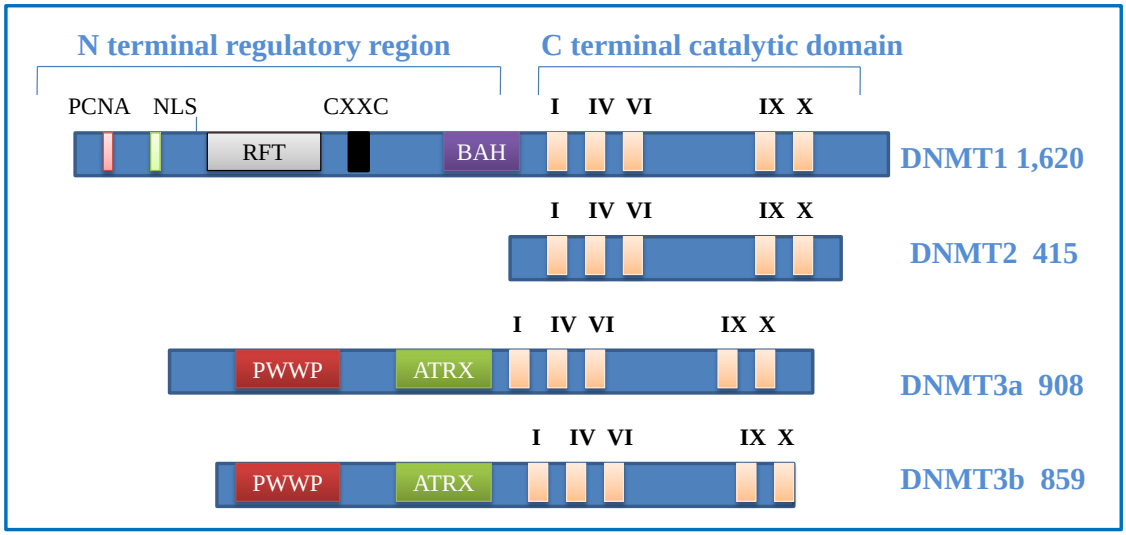


Figure 1-2. Schematic representation of DNMT regulatory and catalytic regions.

Amino terminal regions confer DNMT specificity, whilst the C domain contains highly homologous motifs (X,IX ,VI ,IV ,I). DNMT3A and DNMT3B contain the PWWP domain, implicated in heterochromatin association and the *ATRX* domain composed of *PHD* and C2 - C2 zinc finger, involved in protein-protein interactions. DNMT1 has several specific motifs: BAH, bromo adjacent homology domain for protein/protein interactions, CXXC CpG selective motif, RFT replication foci targeting motif, NLS nuclear localization signal and PCNA interacting domain. Figure adapted from [2, 33].

Besides its general role in mediating protein-protein interactions, the N-terminal domains of some DNMTs are essential for their methyltransferase activity as observed for the DNMT1

enzyme [34]. The catalytic C- terminal domain is highly conserved between DNMT family members and prokaryotic and eukaryotic methyltransferases.

What follows is a brief review of key methyltransferase enzymes involved in methylation establishment and maintenance.

DNMT1- maintenance methyltransferase

DNMT1 was the first identified mammalian methyltransferase. The human DNMT1 gene resides on chromosome 19p13.2, and encodes a 193.5kD protein. It is the most abundantly expressed of all methyltransferases in somatic cells [35]. DNMT1 contains two regions; a diversified N-terminal regulatory region involved in C- domain activity regulation and replication foci targeting, and a conserved C- terminal catalytic domain with homology to bacterial cytosine 5-methylase [36]. Despite this, the C- domain, on its own is inactive and requires the N-terminal domain for its activation [34, 37].

The preferred substrate for this highly processive enzyme is hemi-methylated DNA, research showing that during DNA replication it proccessively methylates only one strand of DNA [38, 39]. DNMT1 recognizes methylated CpGs on the parent DNA strand. The enzyme catalyses transfer of methyl groups from SAM onto the daughter strand ensuring symmetrical methylation patterns [40].

This process is maintained by UHRF1 (ubiquitin like containing PHD and RING finger domains 1, also known as Np95 or ICBP90), a protein responsible for targeting DNMTs to the replication fork [41]. UHRF1 preferentially binds methylated cytosines in hemi-methylated DNA through its methyl binding Su(var)3-9, Enhancer of zeste Trithorax (SET) and RING associated (SRA) domain. However, it can also bind to DNMTs, bringing them into the proximity to the unmethylated cytosines on the daughter strand, which ensures faithful propagation of DNA methylation [42]. UHRF1 also recruits chromatin re-modellers, such as G9a, HDAC1 [43], and histone acetyltransferase Tip60 [44] to aid the establishment of repressive chromatin. Recent crystal structure data has shown UHRF1 as one of the first non- enzymatic DNA binding proteins to use a base flipping mechanism in its interactions with DNA [41].

A role for DNMT1 in transcriptional repression was proposed based on the observed interactions with a variety of proteins such as PCNA, p21WAF1 [45], HDAC1 and HDAC2 [46], MBD2, MBD3 [47] and MeCP2 [48]. There are several tissue - specific splice variants of DNMT1 which differ in their transcription start site and their preference for methylated *versus* unmethylated DNA. DNMT1b is a somatic tissue- specific splice variant. DNMT1o is an oocyte specific isoform with a role in imprinting [49]. DNMT1p, so far observed only in pachytene spermatocytes is believed not to be translated due to inhibition from upstream ORFs [50]. The recently resolved crystal structure of DNMT1 (shown in Figure 2-3) provides considerable insight into the structure and functional regulation of the enzyme [51].

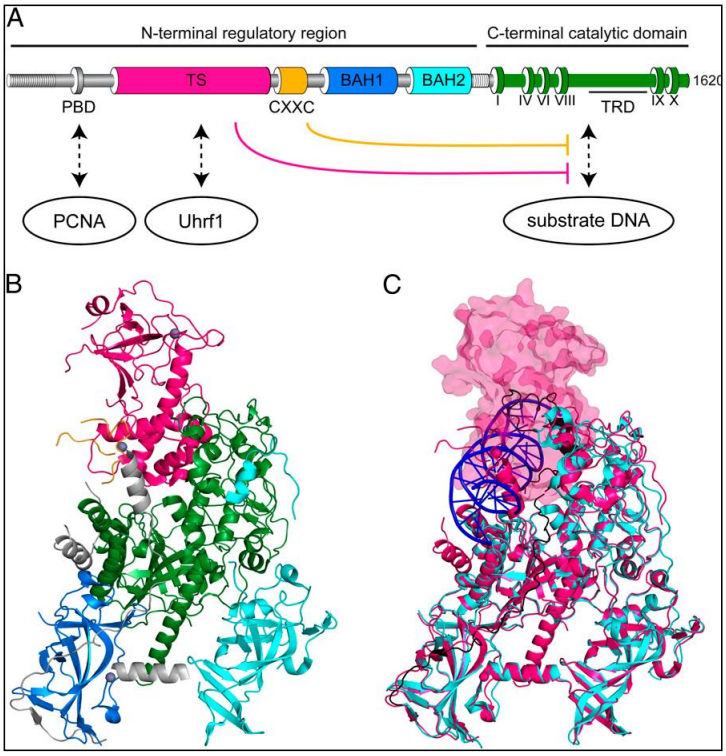


Figure 1-3. Structural representation of DNMT1 and its key interacting partners.

A Functional domains of DNMT1: PBD domain is a site of PCNA replication machinery association and TS domain mediates heterochromatin interactions [52]. It is suggested that TS and CXXC can inhibit substrate DNA binding in an auto-inhibitory function [53]. **B.** Crystal structure of DNMT1. Domains are colour coded as in **A**. **C**. DNA/DNMT1 interaction showing steric clash between CXXC, TS domain and adjacent BAH1linker. Figure from [53].

The above model suggests that the inhibitory domains CXXC, TS and a linker to bromoadjacent homology 1 (BAH1) domain have to be displaced in order to facilitate C- domain DNA binding. The auto-inhibitory effects of the DNMT1 domains fit well with the proposed maintenance activity for this enzyme. Furthermore, this is likely the mechanism that ensures hemi-methylated site

methylation is propagated without errors [54]. The TS domain appears to inhibit constitutive C-domain action, which highlights the requirement for specific activating factors. A proposed candidate for this role is UHRF1, shown to be required for DNMT1 activity *in vivo* [55].

Mouse knockdown studies have shown that DNMT1 silencing leads to abnormalities ranging from aberrant imprinting [56], embryonic lethality [57], and reduced DNA methylation, to re expression of endogenous retroviruses [58].

DNMT1 catalytic activity can be inhibited through its interaction with the pRb protein [59]. This was shown in a study where the over expression of pRb in cells results in reduction of overall genomic 5-methylcytosine levels [60]. DNMT1/pRb interaction at promoters containing E2F transcription factor binding sites is a proposed mechanism for regulating DNA methylation events. It has been postulated that the DNMT1/pRb interaction at precise locations in the genome prevents DNMT1 from binding to PCNA [61]. HDAC1 and HDAC2, two highly homologous histone deactelyase interact with all DNMT family members via their terminal N domains. The exact functional significance of this type of interaction is still unclear although a role in the coordination of re-methylation has been suggested [62].

DNMT2 - TRDMT1

The methyltransferase DNMT2 is located on chromosome 10p15.1. The enzyme is structurally similar to other eukaryotic and prokaryotic DNA methyltransferases. Yet, DNMT2 so far shows only weak *in vivo* and *in vitro* DNA methylation activity [63, 64]. The enzyme was finally proposed to be involved in methylation of small aspartic acid transfer RNA at cytosine 38 in the anticodon loop [65]. Consequently, DNMT2 was renamed tRNA aspartic acid methyltransferase 1 (TRDMT1) in order to reflect better its newly established role. The mechanism employed for RNA methyltransferase activity resembles the catalytic process of the DNMT family, based on the observation that similar residues from C- domain motifs IV, VI, and VIII are involved in catalysis [66]. Interestingly, DNMT2 does not possess the N-terminal domain characteristic of the DNMT family as a whole.

DNMT3 *de novo* methyltransferases

The DNMT3 family of *de novo* methyltransferases has three members; DNMT3A, DNMT3B and DNMT3L. DNMT3A was mapped to human chromosome 2p23, DNMT3B to chromosome 20 q11.2 [35], and DNMT3L is located on chromosome 21q22.3 [67].

The enzyme family appears to be highly conserved throughout evolution suggesting a critical function. Significant sequence homology exists between DNMT3 and DNMT1 family members. DNMT3s do not require hemi-methylated DNA substrate for their activity [68] and are involved in the establishment of *de novo* methylation patterns [69]. However, DNMT3s may also be required for the maintenance function as shown in mouse embryonic stem (ES) cells where there is a level of co-operativity between DNMTs to maintain the methylation of repetitive elements [70, 71]. Recently, it was shown that DNMT3A/B usually co-localize with nucleosomes containing methylated SINE and LINE elements and with CGIs [72]. The DNMT3A/3B PWWP domain helps target the enzyme to pericentromeric heterochromatin, most likely *via* a mechanism that does not involve DNA /protein binding interactions. This is proposed because the PWWP domain of DNMT3A does not bind DNA, and DNMT3B binds non-specifically [73].

The mechanism of action of the two enzymes is different, with DNMT3A acting in a distributive and DNMT3B in a processive manner [74]. This is significant in light of the suggestion that DNMT3B has a preference for sequences such as CpG rich satellite repeats, whilst DNMT3A acts more readily to *de novo* methylate single copy genes [74].

It was demonstrated that DNMT3A represses transcription actively, independent of its C-domain and probably *via* its PHD domain interactions with histone deacetylases (HDAC) and HP1[75, 76].

The DNMT3L enzyme is catalytically inactive. It lacks several key sequences in its conserved region such as a PC active site and an ENV cofactor binding region. Instead, DNMT3L may play a role in targeting active DNMTs to specific locations in the genome. Another proposed role for this enzyme is in scanning highly repetitive DNA regions to direct *de novo* DNA methylation, as demonstrated by observations of silencing in male germ line retrotransposons [77]. DNMT3L, *via* its C-terminal domain, can enhance DNMT3A/B methyltransferase activity *in vitro* [78].

Recently, it was shown that DNMT3L connects DNMT3A2 to nucleosomes in ES cells. The N-terminal cysteine rich region of DNMT3L recognizes unmethylated histone H3K4 tails, and recruits or activates DNMT3A2 to induce *de novo* methylation [79].

DNMT3L expression analysis indicates significant tissue specificity, with DNMT3A predominantly expressed in adult tissue, embryo and ES cells, and DNMT3B in testis, - indicating a role in spermatogenesis [80]. Both DNMT3A and DNMT3B are found to be overexpressed in carcinogenesis contributing to the hypermethylator phenotype in various tissues such as endometrial [81], oesophageal squamous cell carcinoma [82], prostate [83] and breast [84].

Studies on DNMT3A-/-, DNMT3B-/- ES cells have shown loss of *de novo* DNA methylation, loss of global methylation, and more severe embryonic defects compared with each of the single knockouts. These observations have led to the proposal that the functions of DNMT3A and DNMT3B in establishing DNA methylation patterns during developments overlap [85].

Many different isoform variants of DNMT3A/B exist, with a role in a number of epigenetic process ranging from establishment of global methylation patterns during embryonic development [86], gene regulation independent of DNA methylation and modulation of *de novo* methylation [87-89].

Prokaryotic DNA methyltransferases

Bacterial DNA repair mechanisms employ methylation to distinguish between parental and daughter strands and directly repair any replication errors. Methyltransferase enzymes are a component of the restriction- modification system (RM) used by bacteria for protection from foreign DNA, such as bacteriophages. Bacteria use restriction endonucleases to cleave and destroy invading DNA. At the same time, they tag their own DNA with methyl groups to protect it from destruction [90]. Over 3000 such systems are known and these can be subdivided into three categories. Type I RM systems consist of three polypeptide components, each of which is allocated a specific function. ATP dependent interaction of these three polypeptides is required

for restriction and methylation to take place. Best characterized are Type II RM systems and these are most commonly used as experimental tools. The endonuclease and corresponding methyltransferase component act independently of each other and compete for the same DNA recognition sites. Type III RM systems are capable of simultaneous modification and restriction [91]. This is achieved through the interaction of the two components of this system coded by the structural genes *mod* and *res.* Restriction enzyme activity is catalysed by both components, but the *mod* gene product can also function as a modification methylase [92].

Another important role for DNA methylation in bacteria is to control gene expression and cell cycle events. This is achieved in *E.Coli* through the activity of the DNA adenine methylation enzyme (*Dam*) which catalyses methylation of adenine at GATC sites, leading to retention of the origin of replication of *E.Coli*. in the hemi – methylated, inactive state. This prevents further DNA replication events until the cell cycle is completed [93]. Gene expression is regulated through varying methylation patterns at bacterial transcription factor binding sites. *E.Coli* pathogenicity, for example, is influenced by methylation changes at the *Dam* sites in the promoter region of P-pili, which affects its gene expression patterns [94].

Members of the prokaryotic methyltransferase family include M. HpaII (CCGG), M.HhaI (GCGC), M.SssI (CG), M.HaeIII, M.TaqI, and M. RsrI. The bacterial methyltransferases have very similar catalytic activity and the same highly conserved C- domain region as mammalian enzymes [95]. Due to their ability to target and methylate specific DNA sequences, prokaryotic methyltransferases are used extensively in biotechnology applications as important analytical and research tools. The basic mechanism of DNA methylation was to some extent unravelled by studying these enzymes. For example, the important concept of DNA base flipping was first observed from functional studies of M.HhaI. First purified from *Haemophilus haemolyticus*, this Ado-Met dependent methyltransferase methylates the internal cytosine in the tetranucleotide palindrome 5'-GCGC-3'/3'CGCG-5' sequence [96]. The target base is made accessible through the rotation of the sugar phosphate backbone of the specific base out of the DNA helix and into the concave catalytic pocket of the enzyme. This minimizes steric hindrance of the stacked DNA helix [97]. This mechanism of DNA base flipping is now known to be employed by eukaryotic DNMTs and DNA glycosylases [98] to gain access to DNA bases.

1.1.3 DNA demethylation

The mechanistic details of DNA methylation are well understood. Remarkably little is known about the process of active demethylation, and we are still far from a conclusive model for describing the process. Proposed mechanisms include cytosine deamination, 5-methylcytosine DNA glycosylase enzymes, oxidation, and base excision repair enzymes [99]. It appears that the process of DNA demethylation can be either a passive or active one. Passive demethylation following replication excludes DNMT1 from the nucleus [100].

One active DNA methylation model proposes the existence of demethylase enzymes which directly remove the methyl group from ^{5m}C, the subject of which has been controversial and unproven. Recent evidence of cyclical methylation/demethylation in transcriptionally active promoters in mammalian cells further reinforces the idea of an active rapid process [26, 101]. What these studies suggested is that the DNMT3A and DNMT3B enzymes initiate demethylation, and the authors speculate that this is achieved by the enzymes' deaminase activity, based on the evidence that DNMTs can convert C to T and C to U.

Other putative ^{5m}C demethylases have also been put forward as potential candidates for this key epigenetic role, none of which have ever been conclusively confirmed. The first ever reported active demethylase activity was in the nucleoplasm of murine erythroleukemia cells. It demonstrated active removal of tritiated methyl groups from DNA methylated at HpaII sites [102]. This was followed by an observed ribozyme-like demethylase activity in rat myoblast whole-cell extracts [103], which was soon after refuted as an anomaly in experimental procedure [104]. Another disproved demethylase was thymine DNA glycosylase (TDG) put forward as an RNA dependant ^{5m}C glycosylase initially in chick embryos, only to be shown subsequently inefficient at ^{5m}C removal *in vitro* [105, 106]. A role for the methyl binding domain 2 (MBD2) transcriptional repressor which binds methylated DNA both *in vivo* and *in vitro* [107] as a putative demethylase was also proposed [108]. The suggestion that MBD2 can actively and independently remove methyl groups could not be reproduced by other labs. Studies on the nuclear protein GADD45a, and *Xeroderma Pigmentosum* complementation group G (XPG), a DNA repair endonuclease, led to the suggestion that this is an active demethylase. The suggestion was based on the observation that upon overexpression of GADD45a global demethylation could be observed

concurrent with the activation of methylation silenced promoter. Similarly, silencing of GADD45a leads to gene expression silencing and DNA hypermethylation [109]. Gene reactivation and DNA methylation analysis that followed this study were again unable to substantiate the original findings [110].

Recently an entirely new group of enzymes have been implicated in the demethylation process. The Ten Eleven Translocation (TET) family of Fe(II) and 2-oxoglutarate-dependent enzymes and the Activation-Induced (cytidine) Deaminase (AID) protein were initially described as regulators of other processes. TET in leukaemia and AID in antibody diversification were subsequently assigned a new role as the lead candidates in search for the ever elusive demethylase. TET proteins are capable of enzymatically generating 5- hydroxymethylcytosine (5hmeC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) derivatives from the initial methylated cytosines [111, 112]. The overview of the conversion process is shown in Figure 1-4. The 5- hydroxymethylcytosine (5hmeC) was suggested to be an intermediate of the active demethylation pathway given the recent observations of another two cytosine modifications; 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). These modifications were shown to be generated by two step oxidative process involving 5hmeC and TET proteins [113].

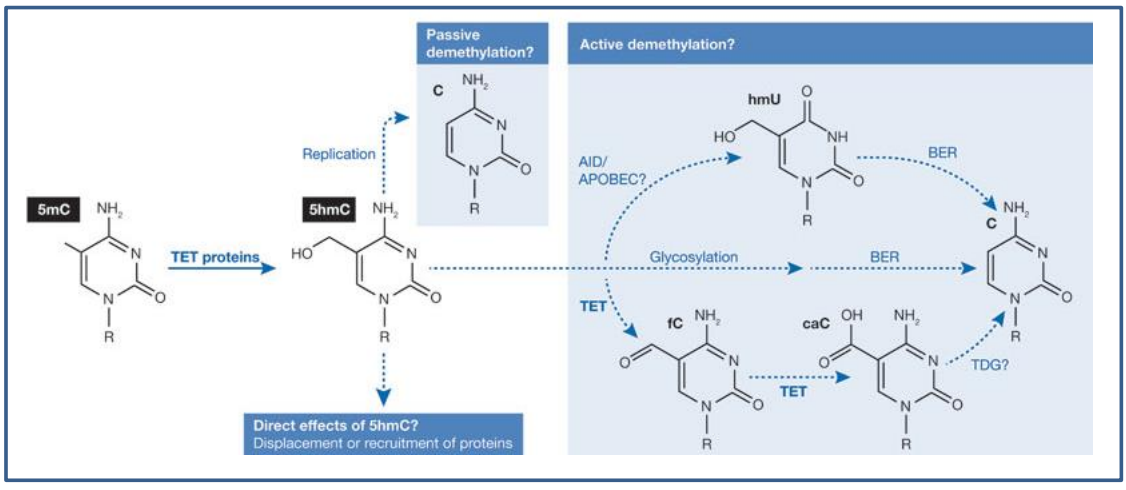


Figure 1-4. Proposed TET mediated active and passive demethylation process.

Proposed mechanism or active and passive demethylation process catalysed by TET proteins. The conversion of 5mC mediated by TET leads to a formation of a 5hmeC intermediate with a putative role in both active and passive demethylation process as well as a direct recruitment or displacement of various proteins. Passive demethylation would probably occur in replication dependant manner, whilst the process of active demethylation would likely involve other 5meC intermediates 5fC and 5caC and DNA repair mechanisms such as deamination mediated by AID/APOBEC, or glycosylation followed by BER. Figure from [114].

Generated by oxidation of 5mC, 5hmeC is found enriched at gene promoters and CGIs in primary cells. It usually correlates with euchromatin and increased levels of transcription [115]. Studies so far indicate that 5hmeC could be a demethylation intermediate, part of a passive mechanism where the presence of 5hmeC impairs re-methylation or DNA repair of modified cytosines leads to a demethylated state [116].

AID/APOBEC family members are another group of enzymes currently pursued for their possible role in demethylation, first demonstrated in zebra fish embryos [117].
The mechanism involves deamination of 5meC to form either 5-methyluracil (5mU) or 5-hydroxymethyluracil. Studies in AID-/- mice show an increase in global hypermethylation [118] but without a profound effect on mice viability, suggesting other compensatory mechanisms are taking place [119]. Apart from the observed zebra fish embryo global demethylation activity, AID was also implicated in loci specific demethylation events. Fusion of mouse ES cells and human fibroblasts led to rapid demethylation at the promoter of two pluripotency genes Oct4 and NANOG followed by transcriptional activation [120]. However, the exact mechanism for its specific action in this case is unclear as AID does not possess a DNA binding domain.

Finally, members of the UDG family of base excision repair (BER) glycosylases such as thymine-DNA glycosylase TDG and single-strand selective mono-functional uracil-DNA glycosylase 1 (SMUG1) could potentially complement existing mechanism by which the above cytosine intermediates are replaced, for example in partnership with TET and AID action [121, 122].

1.1.4 Transcriptional regulation by DNA methylation

Of all the epigenetics processes involved in neoplastic development, promoter region hypermethylation is probably the most prominent. In healthy cells, CGIs are normally unmethylated irrespective of their expression status.

A comprehensive study of methylation changes associated with tumorigenesis in genes which encode tumour suppressors, regulators of apoptosis, DNA repair, and cell cycle has shown that all of the genes that were silenced possessed promoter hypermethylation. Genes studied included tumour suppressors such as *CDKN2A, CDKN2B,APC,BRCA1*, DNA repair genes *MLH1, GSTP1, MGMT* and metastasis and invasion related genes *CDH1, TIMP3* and *DAPK* [123]. These aberrant patterns are frequently cancer or age related. Given that these gene promoters easily succumb to aberrant methylation under given circumstances poses the question as to what keeps a gene promoter methylation free in normal cells?

Possible reasons for the methylation free status of CGIs could be due to the high CG content of the islands preventing normal methyltransferase activity. However, in cancer or X chromosome inactivation such CpGs are obviously methylable. The probable existence of demethylase activities involved in the active removal of methyl CpGs is also a possibility [124]. Finally, it is also likely that TFs binding at CGIs interfere with methyltransferase access/ activity [125]. To better understand the mechanisms behind the occurrence of promoter hypermethylation and protection, we will review some key mechanisms involved in the regulation of an actively transcribing promoter.

1.1.4.1 Gene transcriptional regulation in mammals

Promoters contain a number of specific sequences such as transcription factors binding sites, transcription start site and TATA box recognition elements.

Process of gene transcription commences upon RNA Pol attachment to the promoter via recruitment through both general and specific transcription factors. There are three major classes of RNA Pol enzymes: RNA Pol I involved in ribosomal RNA (rRNA) transcription, RNA Pol II synthesizes mRNA, small nuclear RNAs and microRNAs and RNA Pol III which plays a role in transcribing genes encoding tRNAs and other small RNAs.

The RNA Pol II complex is over 3MDa in size comprising nearly 60 proteins which can be subdivided into three major components.

1) RNA Pol II a 12 polypeptide complex with a molecular weight of over 500 kDa responsible for synthesis and proofreading of RNA [126].
2) Basal transcription factors TFIIA,TFIIB,TFIID,TFIIE,TFIIF and TFIIH that recognize and unwind promoter DNA [127].
3) A 25 subunit large polypeptide component with a molecular weight of 1MDa called a Mediator which interacts with activators and repressors [128].

The process of transcriptional activation can best be explained through a step-wise model which proposes chromatin opening [129], followed by binding of transcriptional activators to one or more regulatory sequences located in the promoter region of a gene. A key feature of this process is the assembly of the pre-initiation complex (PIC) mediated by basal transcription factors TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH.

Initially, the TBP (TATA box binding protein) subunit of TFIID in co-operation with TAFs (TBP associated factors) binds to the TATA box.

TFIIB (transcription factor II B) then joins this complex, recruiting RNA Pol II which together with TFIIE, TFIIF and TFIIH leads to the promoter associated formation of PIC. This assembled complex forms a mobile clamp around promoter DNA with the active site of the enzyme buried in the cleft between the two largest subunits RBP1 and RBP2. ATP dependent TFIIH helicase mediated promoter melting occurs leading to transcriptional initiation [130]. In the RNA Pol II complex the DNA is further unwound and the DNA template is transcribed into a complementary RNA transcript where new nucleotides get continuously added to the 3'OH end of the RNA [131]. Finally, promoter clearance and transcriptional elongation are facilitated through phosphorylation of the RNA Pol II subunit Rpb1- CTD (C-terminal domain) by TFIIH [132].

1.1.4.2 DNA methylation mediated transcriptional regulation in mammals

Accumulated evidence so far suggests that promoter methylation is associated with the stable silencing. There are a number of proposed mechanisms by which this occurs. The first is that the presence of methylated cytosines directly blocks and thereby prevents methylation sensitive

transcription factors such as c-Myb, YY1, and MLTF from binding to their recognition sites in EMSA assays. The transcription factor Sp1, although insensitive to CpG methylation within its recognition site is sensitive to combination of CpG and non-CpG methylation in the same site, suggesting methylation mediated regulation of Sp1 in stem cell scenario. Insensitivity to CpG methylation has led to the proposal that Sp1 binding serves as a mechanism for maintenance of methylation - free CpG islands [133].

The second proposed mechanism involves the binding of methyl cytosine binding domain (MBD*)* proteins, which in their function as "readers" of the epigenetic code can selectively bind to methylated DNA to facilitate transcriptional silence. There are five known members of the MBD family with MBD1, MBD2, MBD3 and MeCP2 which share a methyl - CpG binding domain with a high affinity for methylated DNA and are partially involved in transcriptional repression [134]. In contrast, MBD4 with its DNA N-glycosylase activity is believed to be involved in DNA repair [135].

The MBD protein family members "read" methylation patterns and translate them into an appropriate chromatin functional state. This is achieved through interactions with various nuclear factors [136]. MeCP2, MBD1, MBD2 and Kaiso protein when recruited to ^{me}C interact with co-repressor complexes such as NurD/HDAC leading to an inactive chromatin configuration and gene silencing [134, 137]. MeCP2 is involved in transcriptional regulation through its transcriptional regulation domain (TRD) interactions with chromatin modifying enzymes such as, histone deacetylase corepressor Sin3a [138] or the SWI/SNF chromatin re-modelling catalytic component Brahma (Brm) [139].

1.1.5 Chromatin modifications

Chromatin is a dynamic macro molecular structure. It consists of DNA nucleosome repeats that can be classified into euchromatin and heterochromatin. Euchromatin is a decondensed form that permits transcription, whereas heterochromatin is highly compacted and transcriptionally silent. Heterochromatin can be subdivided into constitutive - permanently silent and facultative - repressed states [140].

Histones mediate DNA organization into nucleosomes. A nucleosome consists of 147bp DNA wrapped in left handed superhelical turns around an octamer of core histones H2A, H2B, H3

and H4 [141]. Subsequent nucleosomes are connected via linker DNA of characteristic length. One histone linker molecule H1 binds in the region which is at the point where DNA enters and exits the nucleosome [142]. Nucleosomes are further stabilized by H1: linker DNA interactions. DNA : histone octamer binding is through interaction with the phosphodiester backbone *via* three to six hydrogen bonds, which permits DNA sequence independent binding to the core of the nucleosome [143]. Tight aggregation of nucleosomes generates a transcriptionally repressed heterochromatin structure.

1.1.5.1 Histone modifications

Transcriptional activity is thought to be influenced by covalent modifications to the histone tail residues, which protrude out of the nucleosome. These modifications include phosphorylation of serine and threonine residues, acetylation of lysine residues, methylation of lysine and arginine, ubiquitination, sumoylation, ribosylation, glycosylation, biotinylation and carbonylation [144]. The modifications can mark regions of chromosome for specialized functions.

By default, heterochromatin is non-permissive to transcription. Highly repressed chromatin structure arises from DNA: histone and DNA: histone tail interactions.

Histone modifications are established through the activity of chromatin modifying enzymes. The four principal enzymatic systems involved in histone modification catalysis and their antagonists, that are involved in modification reversal, are listed in Table 1.

Histone modifying enzyme	**Histone modification**	**Antagonist**
HAT –Histone acetylase	Lysine acetylation	HDAC- Histone deacetylase
Serine,Threonine Kinase	Phosphorylation	PPtase- phosphatases
PRMT- Protein arginine methyltransferase	Arginine methylation	Arginine Deiminase
HKMT-Histone lysine methyltransferase	Lysine methylation	LSD1-Lysine specific demethylase (amine oxidase)

Table 1-1. Histone modifying enzymes, modifications and their antagonists.

Most known histone modifications correlate with either activating or repressive activity in the cells, although splicing and replication have also been inferred. The functional consequences of these histone modifications are disruption of chromatin architecture which can affect DNA: histone interactions.Table 1-2 shows a list of best characterized histone modifications. The information was gathered by HIstome: The histone infobase [145].

Modification	**Histone position**
Lysine Acetylation	H3: K4, K9, K14, K18, K23, K27,K36,K56 H4: K5, K8, K12, K16, K91 H2B: K5, K12, K15, K16, K20, K46, K120 H1K25, H2AK5,
Serine,Threonine,Tyrosine Phosphorylation	H1: S171, S172, S17, S186, S26, T10, T137, T145, T153, T154,T17, T30 H3: S10, S28, S31, S6,T11,T3, T45, T6, Y41, H4S1
Arginine methylation	H3: $R17^{me1}$, $R17^{me2}$, R26, $R2^{me1}$, $R2^{me2}$, $R8^{me2}$, H4R3, $H4R3^{me2}$
Lysine methylation	H3: $K27^{me1,me2,me3}$, $K36^{me1,me2,me3}$, $K4^{me1,me2,me3}$, $K79^{me1,me2,me3}$,$K9^{me1,me2,me3}$, $H420^{me1,me2,me3}$
Lysine biotinylation	H2A: K125,127,129,K13,K9 H3: K18, K19, K9, H4K12, H4K8
Lysine ribosylation	H2AK13, H2BK30, H3K27, H3K37, H4K16

Lysine ubiquitination	H4K91, H2AK119, H2AK121, H2BK120

For instance, HAT catalysed histone acetylation of N-tail lysine residue results in loss of positive charge which can potentially weaken DNA: histone interaction [146]. This leads to a transcriptionally permissive, open chromatin configuration. Similarly, HDAC mediated deacetylation of histones restores the lysine positive charge, promoting a strong histone: DNA phosphate backbone interaction resulting in formation of a condensed transcriptionally repressive chromatin signature.

Gene regulation by histone modifications is further regulated through interactions with non-histone proteins and chromatin re-modellers. Regulated genes in their repressed state have promoters with nucleosomes and TFs effectively competing for the regulatory sequences. These promoters rely on chromatin re-modellers for their activation [147].

In humans, there are three classes of ATP dependant re-modellers; the hSWI/SNF, ISWFI families and NuRD. Studies indicate that such re-modellers can expose nucleosomal DNA either by sliding of DNA in relation to the histone octamer [148] or by conformational changes, which lead to an alternate nucleosomal DNA exposure on the surface of the histone octamer. ISWI re-modellers with the exception of NURF and Isw1b help organize nucleosomes promoting gene repression[149]. The ISWI/NURF chromatin remodelling complex is involved in nucleosome positioning, usually by sliding nucleosomes along the DNA [150]. In *Drosophila,* NURF catalyses progressive distribution of nucleosomes along the *hsp70* promoter DNA, whilst maintaining the histone octamer integrity [151]. This is achieved through the use of DNA linkers to facilitate nucleosome sliding, to create uniform nucleosome arrays. SWI/SNF family re-modellers on the other hand, disorganize chromatin by nucleosome ejection leading to gene activation. It is believed that they are targeted to active promoters due to their ability to bind acetylated histone tails [152, 153]. SWI/SNF or Brahma complexes usually work in synergy with chromatin modifying enzymes and use ATP-hydrolysis to influence nucleosomal remodelling and disrupt histone : DNA contacts [154].

The binding of non-histone proteins to modified histone residues can recruit further histone modifying enzymes. For instance, the Heterochromatin protein 1 (HP1) via its chromo domain reads H3K9 trimethylation to transcriptionally regulate surrounding chromatin [155]. HP1 has three isoforms, which can be further modified. HP1alpha, and HP1beta associate with heterochromatin region and serve as guardians of methyl H3K9 mediated silencing, but phosphorylation of HP1gamma at Ser 83 is typically associated with euchromatin and transcriptional elongation [156].

As shown in Table 1 two groups of enzymes, HATs and their antagonist HADACs, catalyse the processes of histone acetylation and deacetylation respectively. Core histone lysine residues in the N-terminus are hyperacetylated in transcriptionally active genes [157], and the overall process of addition and removal of acetyl groups ifs facilitated by HAT (histone acetyl transferase) [158] and HDAC (histone deacetylase) [159] enzyme complexes respectively. These complexes, when targeted to gene promoter regions, create specific patterns of acetylation which influence transcription. HATs are brought to promoters by interaction with DNA bound activators resulting in the recruitment of the general transcription machinery [160]. Histone tail acetylation occurs in the presence of the co-factor acetyl Coenzyme A. Acetylation occurs predominantly on the lysine residues of histones H3 and H4. Transfer of an acetyl group results in neutralization of positive charge, and a subsequent reduced electrostatic interaction with DNA. This opens up the chromatin structure [161]. The lowered positive charge on acetylated N termini which affects the stability of DNA interactions might also be a contributing factor to TF recruitment. HATs can exist as Type A (nuclear) and Type B (cytoplasmic) involved in free histone acetylation. It has been shown that histone acetylation can also reduce nucleosome compaction by disrupting histone tail internucleosomal interactions [162].

Nuclear HATs include diverse families of Gcn5/PCAF, MYST and p300/CPB[146]. Of these, the best studied is p300/CBP, capable of non -histone protein acetylation, for instance the c-Myc oncoprotein [163]. Furthermore, the same HAT protein was shown to have a role in cell cycle control G1/S progression via its interaction with retinoblastoma RB [164].

HDAC activity mediates removal of acetyl groups from N-acetyl lysine and is associated with transcriptional repression. The enzymes are part of a superfamily and can be divided into four classes (I to IV), which are then further subdivided to give 18 different members (HDAC 1-11)

and Sirtuin family members (SIRT 1-7)[165]. In the yeast model, permanently silent heterochromatin was shown to be formed through the interactions of histone H3 and H4 tails with the SIR3 and SIR4 proteins, which prevents tail acetylation [166]. Apart from Class III, all of the HDACs are Zinc dependent enzymes [167].

Histone methylation can also be associated with heterochromatic DNA silencing as demonstrated for the lysine methyltransferase enzyme Su(var)39 (Suppressor of variegation 39), which methylates lysine 9 of histone H3. This modification was shown to function in heterochromatic silencing in *D. melanogaster* [168].

An additional level of complexity is afforded by the fact that lysines can be mono, di or tri - methylated, whereas arginines can only be mono or di-methylated [169, 170]. The best characterized lysine methylated sites are on histone H3 (K4, K9, K27, K36, K79) and histone H4(K20). Marks associated with transcriptional activation are histone H3K4, H3K36 and H3K79, whilst all the others are repressive marks [171].

Histone methylation is regulated by methyltransferases which aid transfer of methyl groups from the SAM donor site to either lysine's ε-amino or arginine ω-guanidino group.

HKMT-Histone lysine methyltransferase activity is mainly facilitated through the catalytic SET domain, in a site specific and a level of methylation specific manner (i.e. mono, di or tri methylation. The exception is Dot 1, H3K79 specific enzyme [146].

Prominent members of the SET domain super family of proteins include Suppressor of variegation 3-9 (SUV39), Polycomb-group chromatin regulator Enhancer of Zeste (EZ), SET1, SET2, EZ, RIZ, SMYD, and SUV4-20 families. The SUV39 family specifically methylates H3K9 residues [172]. Members of the family are SUV39H1, SUV39H2, G9a, GLP1, ESET, and CLLL8 and they can be associated with both heterochromatin and euchromatin structure [173]. For instance, HP1 protein was shown to associate with SUV39H1 giving rise to a repressive chromatin architecture [174].

The EZ subgroup of HKMTs belongs to Polycomb repressive complex (PcG), involved in gene silencing. Group members EZH1 and EZH2 mediate methylation of H3K27 as a part of PRC/EED-EZH1/2 complex [175]. H3K27/H3K4me3 is often seen as a bivalent switch, marking

a region that is poised to adopt either an active or repressed state [176]. Genome wide studies indicate that H3K27me3 typically localizes with active promoters, throughout gene bodies, and in intergenic regions whilst H3K4me3 can be found at promoters and the 5′ ends of genes, together with histone acetylation. Recently transcribed genes are marked by H3K36me3, all of which suggests that specific modifications serve to index structurally and functionally distinct chromatin regions [177]. Arginine methylation of histone tails is catalysed by the PRMT family of proteins. They can be subdivided in five Type I enzymes (PRMT1, 3, 4, 6 & 8) and a single Type II enzyme (PRMT5) [169].

Initially believed to be an irreversible process, histone methylation is now known to be dynamic. This was experimentally confirmed with the discovery of the first histone demethylase the lysine specific demethylase, LSD1 [178, 179]. A explosion in the field identified a number of additional demethylases, which can now be grouped into three key classes; LSD1, Jumonji C domain containing histone demethylase (JHDM), and peptidylaragnine demethylase 4 (PADI4). LSD1 removes methyl groups from mono and di methylated H3K4 and H3K9 residues [179, 180]. Nucleosomal recognition by LSD1 is conferred through its interaction with co-repressor complexes such as CoREST [181].

The JHDM family of histone demethylases are capable of demethylating tri-methylated lysines and possess the catalytic jumonji domain [182, 183]. JHDM1 preferentially demethylates H3K36me1/me2; JHDM2 demethylates trimethylated H3K9 and H3K36, whilst JHDM3 shows preference for H3K36me3 [184].

Histone phosphorylation is a process modulated by kinases and phosphatases and involves the addition or removal of a phosphate group from ATP onto the hydroxyl group of serine, threonine or tyrosine side chains [185]. Serine 10 histone H3 phosphorylation has been linked to transcriptional activation [186]. The presence of negatively charged phosphate groups along the histone tail is believed to reduce their affinity for DNA [187]. It has also been found that several HATs have an increased affinity for phosphorylated serine on histone H3 [188]. It is important to note however that histone phosphorylation has also been shown to be involved in chromosome compaction during mitosis and meiosis, for example, H3 phosphorylation on serine 10 and serine 28 [189] and expansion during gene expression. Both phosphorylation and

acetylation are required for transcriptional activation, with phosphorylation promoting acetylation [190].

Another layer of chromatin mediated transcriptional modulation comes from histone variant substitutions. These substitutions appear to index chromosomes for a specialized function. For example, CENP-A (histone H3 centromere specific variant) is essential for chromosome segregation [191]. The whole process of histone variant replacement is not restricted to any particular phase of the cell cycle but occurs in response to on-going mechanisms. For example, the histone H2 variant, H2.A.X. responds to DNA damage lesions and recruits repair complexes to the damaged sites [192]. Gene activation can also be influenced by the presence of specific histone variants such as H2A.Z. This is believed to form a less stable nucleosome, allowing nucleosome eviction and exposure of binding sites within a promoter [193]. H2A.Z is highly enriched at open TATA-less promoters [194]. This histone variant has a unique N-terminal tail which becomes acetylated when the gene is active. Histone variant H3.3 was shown to have a role in heterochromatin formation and silencing *via* RNA-interference (RNAi) pathway during reprogramming [195].

Modifications to the N-tail of histones have been extensively analysed. However, the core structure of histones can also be the subject of modifications, with knock on effects on the overall nucleosomal structure. An example of such modification was the observed phosphorylation of H3T45 mediated by kinase C-delta enzyme in neutrophil apoptotic cells [196].

In summary, histone modifications exert their influence on chromatin architecture either directly or by affecting effector molecule interactions. DNA methylation of CpG residues typically associates with transcriptional silencing, while strong evidence proposes a mechanistical link between DNA methylation and histone modifications. A number of interactions coupling DNA methylation machinery components with histone modifiers include methyl-CpG binding proteins, DNMT1 and DNMT3B which were all shown capable of HDAC activity recruitment ultimately resulting in chromatin remodelling and transcriptional silencing [197-199]. The Polycomb group repressive complex component EZH2 was also shown to associate with DNMT activity *in vivo* and required for DNA methylation of its target promoters [200].

1.2 Myelodysplastic Syndrome

The Myelodysplastic Syndrome (MDS) is a heterogeneous group of a hematopoietic clonal stem cell disorders, manifested in progressive bone marrow failure, which often transforms to Acute Myeloid Leukaemia (AML). A common clinical hallmark of the disease is defective haematopoiesis, which affects erythrocytic, granulocytic and megakaryocytic lineages. The effects are defective peripheral blood cell production (cytopenia), as well as hypercellular "dysplastic looking" bone marrow [201]. MDS is characterized by cytogenetic abnormalities such as deletions of chromosome 5 and/or 7 in patients with poor prognosis, isolated deletion of 5q or trisomy 8 that define different patient subgroups [202]. The epigenetic components of DNA methylation, histone modifications and microRNAs are key drivers of the disease phenotype in MDS [203].

1.2.1 Epigenetic factors in MDS

The role of aberrant DNA methylation and histone modifications in altering neoplastic cell physiology has been a subject of intense interest over the last decade or so. Cancer is marked by several key DNA methylation events. Firstly there is a global loss of methylation from transcriptionally silenced loci such as repetitive elements, satellite DNA and intron located CpGs [204]. Increased DNMT expression in neoplastic transformations has also been observed for both DNMT1 and DNMT3s [205]. Finally, unmethylated CGIs become hypermethylated, probably as a consequence of chronic exposure to DNMTs or through a loss/disruption of promoter protective mechanisms[206]. Gene silencing process associated with promoter hypermethylation has been described for a number of genes, notably *RB1*, *p16*, *VHL* and *MLH1* [207-210].

Analysis of genome - wide DNA methylation in a large cohort of MDS patients has shown a change in global methylation patterns compared with age matched healthy individuals. Although, this review focuses on one particular gene, *CDKN2B* that is of interest to this project, it should be highlighted that promoter hypermethylation has also been observed in several

other genes, in MDS. These genes are *Hypermethylated In Cancer* (*HIC*), *E cadherin* [211], and *Oestrogen Receptor* (*ER*). The functional consequences of their transcriptional silencing are unknown. 65% of MDS patients exhibit *Calcitonin* gene hypermethylation [212], and 31% hypermethylation of *Suppressor Of Cytokines (SOCS) SOCS-1* gene [213].

The promoter region of the tumour suppressor *CDKN2B* gene is frequently hypermethylated in MDS [214]. There is also a correlation between *CDKN2B* methylation and DNMTs over expression [215] in samples from patients with AML. In contrast, this does not occur in MDS [216]. In normal haematopoiesis *CDKN2B* is responsible for maintaining a hematopoietic stem cells in a quiescent state [217]. This gene is upregulated during granulocytic and megakaryocytic differentiation and down regulated at erythroid commitment [218]. In MDS, *CDKN2B* silencing is proposed to result in increased proliferation of stem cells in the absence of differentiation leading to cytopenia and excess blasts [219, 220].

Methylation profiling studies have shown that it is possible to use this gene as a prognostic and diagnostic marker for MDS. *CDKN2B* promoter methylation status analysis in patients with MDS, prior to 5-Azacytidine treatment showed promoter hypermethylation in 68%, $CD34^{+}$, 77% $CD33^{+}$, 86% $CD34^{-}/CD33^{-}$ cells using MS-PCR. Following 5-Azacytidine treatment no significant changes in *CDKN2B* methylation pattern could be observed [221]. It has been suggested that the incidence of *CDKN2B* promoter methylation can be correlated with blast expansion and disease progression towards AML, but this is yet to be fully confirmed [222]. At the early stages of MDS, levels of *CDKN2B* promoter methylation are relatively low (23%),but increase with the disease progression [223]. Gene inactivation and observed *CDKN2B* promoter hypermethylation is associated with high risk MDS, with incidence ranging from 58% for Chronic Myelomonocytic Leukemia (CMML) [224] to 60-75% in MDS/AML [225]. CpG methylation density and frequency in therapy - related MDS and AML was shown to increase disease progression and severity and is prevalent in patients with 7q deletions [226]. Histone modifications have also been implicated in carcinogenesis related changes in gene expression patterns. Histone modifying enzymes were observed to be genetically altered, translocated, silenced or overexpressed in various malignancies.

Examples of such events include acetyltransferase CBP mutation and H3K4 methylase MLL re-arrangement in acute leukaemia [227, 228]. An activating mutation of Janus kinase 2 (JAK2) tyrosine kinase was shown to phosphorylate histone H3Y41 thereby preventing its interaction with HP1α in haematological malignancies [229]. Another H3K9 methylase RIZ-1 was shown to be downregulated in Chronic Myelogenous Leukaemia (CML) patients [230] and in prostate cancer overexpression of H3K27 tri-methylase, EZH2, directly co-relates with the disease progression [231].

1.2.2 Methyltransferase Inhibitor therapy

DNA methyltransferase inhibitor (MTI) drugs are currently being investigated for their potential therapeutic applications in the treatment of cancer. The observation that *CDKN2B* promoter hypermethylation, is potentially reversible has led to attempts to use MTIs as a treatment alternative in MDS [222]. Treatment of MDS patients with MTIs such as 5- Azacytidine (5-Aza) and its nucleoside analogue 5 Aza 2' deoxycytidine was shown to be effective in Phase III trials in up to 60% of patients [222]. 5- Azacytidine is a cytotoxic nucleoside chemical compound first synthesized in 1964. It was shown to incorporate into DNA and bind to *DNMTs* irreversibly [232]. Furthermore, 5-Aza leads to genome - wide DNA demethylation and inhibition of *de novo* methylation [233]. Three Cancer and Leukemia Group B (CALGB) and a Phase III randomised study have been conducted so far to assess the effectiveness of 5-Aza in MDS treatment. These studies indicate an overall response rate of around 50%, with 12% achieving complete remission, 15-25% partial remission and 12%-27% an improvement in their blood counts. All studies noted the improvements to the patients quality of life, survival rates and time to remission [222].

1.3 *CDKN2B* – Role in the Cell Cycle

The mitotic cell cycle involves the duplication of chromosomes and other cellular components and their subsequent segregation into each of two daughter cells. The fidelity of the process, its timing, and coordination, is tightly regulated by a cell cycle control feedback loop system. Cyclin - Cdk complexes are extremely important in these processes (Figure 1-5). The loop consists of several significant stages and checkpoints: Gap 0 (G0) cell resting phase, Gap 1 (G1), first checkpoint, Synthesis (S), DNA replication phase, Gap 2 (G2) second checkpoint, and Mitosis (M), cell division stage.

1.3.1 Cell cycle control system

The cyclin dependent kinases (Cdk) are a family of serine/threonine kinases (34-40kDa) involved in cell cycle control. There are twelve known human Cdks all showing characteristic motifs [234]. These include a T – activation loop, the PSTAIRE and L12 helix, the highly conserved 100 amino acid Cyclin Box, required for Cdk binding and activation, and the Degradation Box located in the N – terminal region, which confers specificity to different classes of cyclins. The process of Cdk activation involves binding of ATP at the Cdk active site which is usually shielded by the T loop, followed by binding of a cyclin at the cyclin box, which results in conformational changes that facilitate phosphotransfer from ATP to the neighbouring threonine residue, catalyzed by Cdk-activating kinases CAK [235]. The levels of cyclin proteins were found to oscillate as the cell cycle progresses [236]. Transition through the early stages of the cycle is associated with activation by Cdk6-cyclin D3 and Cdk4-cyclin D2/3 complexes. Cells are prevented from entering cell cycle quiescence by inhibition of these complexes by the INK4 family of proteins ($p16^{INK4A}$, $p15^{INK4B}$, $p18^{INK4C}$, $p19^{INK4D}$) and Cip/Kip proteins ($p21^{Cip1}$, $p27^{Kip1}$ and $p57^{Kip2}$). Two other pertinent factors in G0/ G1 progression are the pRB and p130 proteins. Initially hypo - phosphorylated in quiescent cells, these proteins are capable of restricting transcription through binding to various transcription factors such as E2Fs and by recruiting HDAC and other chromatin modifiers. Phosphorylation of pRb and p130 by Cyclin D-cdk4/6 and

Cyclin E-Cdk2 to hyper – phosphorylated states releases E2F transcription factors from repression. E2Fs then induce genes that are required for entry into S - phase and transition through the remainder of the cell cycle. Cyclin E –cdk2 complexes are rapidly degraded by the anaphase promoting complex (APC) as the cell enters S phase followed by the Cyclin A-cdk2 activity through the proliferating cell nuclear antigen (PCNA) interaction. These interactions lead to DNA replication initiation, followed by inhibition of new replication complexes thereby preventing DNA replication initiation until the next cycle [237].

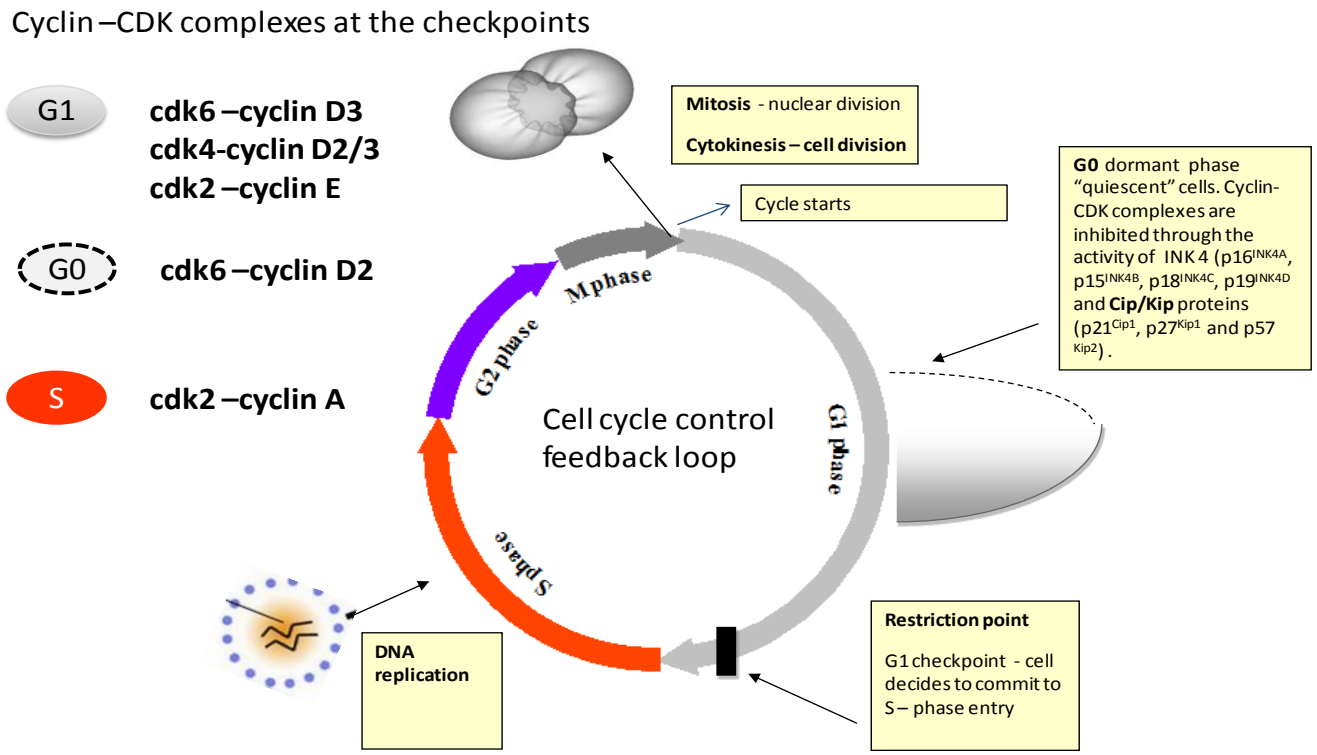


Figure 1-5. The cell cycle control checkpoints. The cell cycle is characterized by DNA replication and chromosome duplication is S phase, followed by M phase, which is composed of mitosis and cytokinesis. The two phases are separated by gap phase G1 before S phase and G2 before M phase. These two phases allow the cell to respond to external stimuli or internal DNA damage. During G1, the cell becomes committed to enter S – phase by passing the Restriction Point. G0 phase is a prolonged nondividing stage. Different Cdk/Cyclin complexes govern different phases of cell cycle transitions, where each cyclin-cdk complex acts on the next step in the sequence, ensuring orderly progression through the cell cycle stages. Figure adapted from [238].

1.3.2 Cdk inhibitors

There are two distinct groups of cdk inhibitor proteins, which can be grouped into INK4 and Cip/Kip families. The Inhibitor of cyclin dependent kinase 4 (INK4) family specifically inhibit cdk4 and cdk6 [239]. They share a common 32 amino acid ankyrin motif, a protein-protein interaction domain which enables each member of the family to bind to cdk [240]. $p16^{INK4A}$ encoded by *CDKN2A* shares 82% sequence homology with $p15^{INK4A}$ which encodes *CDKN2B*.

The expression of *CDKN2B* is stimulated by Miz-1 regulatory protein during quiescence. Miz-1 binds to the Initiator response (Inr) region of *CDKN2B*. It then interacts with the Myc/Max complex, which is also recruited to the *CDKN2B* promoter region to inhibit its transcriptional activation [241]. *CDKN2B* expression was observed to increase dramatically upon stimulation with anti-mitogen transforming growth factor beta (TGF-β) [242]. The process is mediated by the SMAD group of proteins. SMAD activity is directed at Miz-1 either directly, or through E2F4 and 5 mediated c-Myc inhibition [243]. Loss of *CDKN2B* function arsing as a consequence of DNA deletions or mutations results in increased proliferation. This was shown in studies on a hematopoietic CD34+ pluripotent progenitor stem cells, consistent with a role for *CDKN2B* in maintaining quiescence [244].

Other members of the INK4 family are $p18^{INK4C}$ encoded by *CDKN2C* and $p19^{INK4D}$ encoded by *CDKN2D*. *CDKN2C* is required for T cell quiescence [245] and *CDKN2D* may be required in proliferating macrophages during S – phase [246]. The INK4 group of proteins binds cdk4 and cdk6 that are not bound by cyclins, whereas Cip/Kip family members bind cdk-cyclin complexes [247].

$p21^{Cip1}$ is encoded by *CDKN1A*. The gene is induced during the G0-G1 transition and acts as a chaperone that enhances cdk-cyclin stability. It can also be induced at higher levels for example by p53, leading to cell cycle arrest at G1 and G2/M. This protein acts as a cyclin dependent kinase inhibitor of cyclin-cdk2 and Cyclin –cdk4 complexes [248]. Several other roles for *CDKN1A* have been proposed in DNA damage responses such as S phase DNA replication via PCNA interactions [249], and apoptosis [250]. $p27^{Kip1}$ encoded by *CDKN1B* has a role in regulating the transition from G1 to S. In quiescent cells it binds and inhibits cyclin E-cdk2. In

response to mitogenic stimulation, the cdk – cyclin complexes are re –sorted; *CDKN1B* is phosphorylated by one of the three kinases LYN, BCR-ABL or SRC [251, 252] , polyubiquilated by the SCF complex, and degraded by the proteosomes pathway [253, 254].

1.4 Project Description

This project will study the epigenetic and functional consequences resulting from variable targeted DNA methylation of the *CDKN2B* gene promoter. The rationale for initially targeting specific promoter regions is based on transcription factor binding site information. Previous work in our lab has highlighted methylation patterns in MDS patients where just the CpGs present within the c-Myb or YY1 sites are methylated, suggesting such methylation as a potential early event in *CDKN2B* gene shutdown. Zinc finger methyltransferase fusions have been designed to target these regions. Synthetic, *CDKN2B* promoter - specific zinc finger modules fused with prokaryotic methyltransferase have been used to determine the role methylation plays in regulating transcription processes in the context of aberrant methylation in disease. Previous studies in our department showed a variable methylation of *CDKN2B* promoter in MDS. The promoter methylation study in myeloid, lymphoid and CD34+ stem/progenitor cells obtained from MDS patient samples showed a low level, non-clonal and variegated methylation patterns (Figure 1-6). The study further confirmed previously reported intra- and inter-individual heterogeneity in *CDKN2B* methylation in MDS [255, 256].

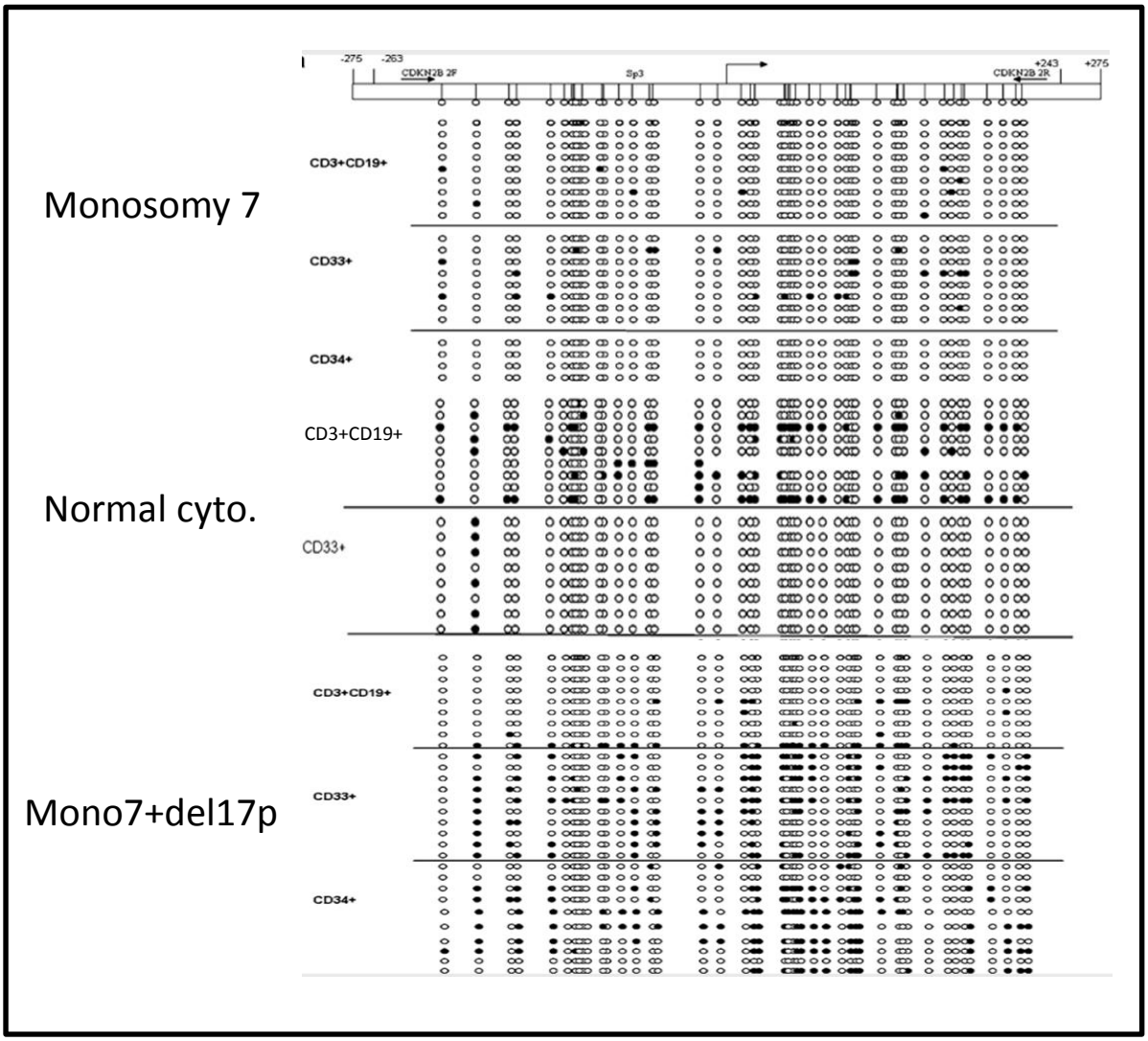


Figure 1-6. The variable methylation patterns at the *CDKN2B* promoter in MDS

This figure originates from [256] and shows *CDKN2B* promoter methylation analysis of MDS patients samples. Patients presented with chromosome 7 loss, with and without deletion of 17p or normal cytogenetics. The myeloid ($CD33^+$), lymphoid ($CD3^+$, $CD19^+$) and $CD34^+$ stem/progenitor cells were isolated from the bone marrow of the 17 patients and analysed for their *CDKN2B* methylation status by bisulfite genomic sequencing (BGS).

1.4.1 Zinc finger module platform technology

Classical zinc finger proteins (ZFP) $Cys_2His_2(C_2H_2)$ were first discovered in the transcription factor TFIIIA and are now known to exist in 2% of all human transcription factors [257]. The zinc finger domain typically has two antiparallel beta strands and an alpha helix (ββα) structure of around 30 amino acids in size, stabilized by the presence of a zinc ion bound by two cysteine residues in the β sheets, and two histidine residues in the α helix. Alpha helix residues contain DNA interacting regions with key amino acids (circled and numbered relative to their helical positions in Figure 1-7), capable of contacting three base pairs of DNA. The specificity of the zinc finger can be manipulated through the variation in the way the helix is presented. As a consequence, the zinc finger motif has been widely used as a template for the rational design and construction of DNA sequence specific proteins, due to its ability to bind to virtually any DNA sequence.

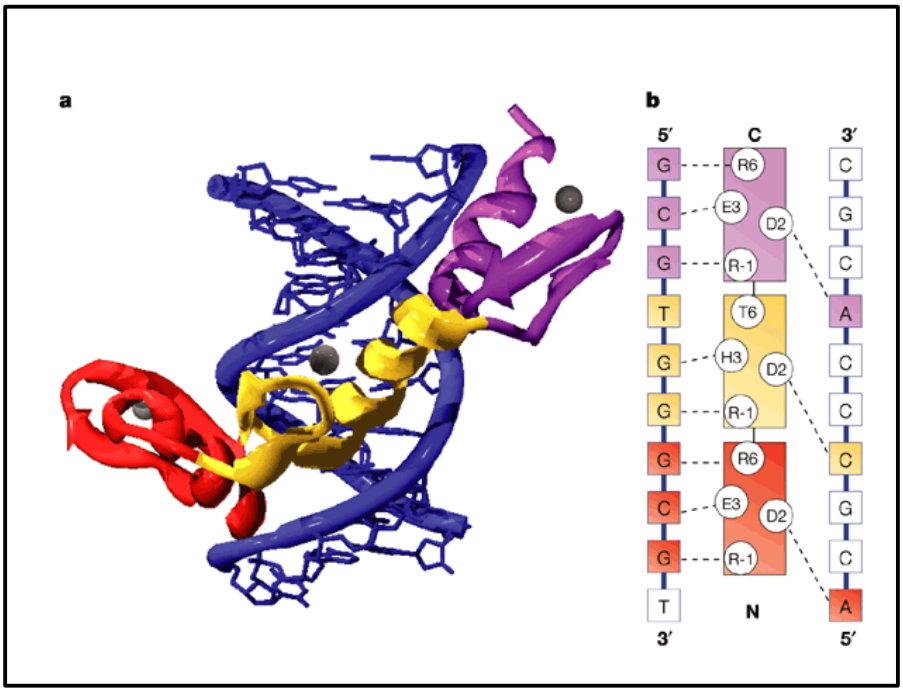


Figure 1-7. The Zif268- DNA showing key residues involved.

A DNA helix (in blue) interacting with a three module ZF. Each domain is coloured differently. Zinc ions which help to stabilize the structure are shown in grey. There is more than one zinc ion shown probably an artefact from x-ray crystallography. B.The key residues involved in making contact are colour coded -1,2,3 and 6 relative to the first amino acid residue of the helix have a major impact on the sequence specificity. Other residues in the vicinity also play an important role in DNA binding interactions. Reproduced from [258].

Much of the information on zinc finger DNA interactions comes from studies on the three zinc finger protein Zif268 [259]. Zif268 has been one of the most frequently used zinc fingers in a variety of design, modelling and DNA targeting studies due to its well -known and characterized structure [260]. Initially discovered in mouse, Zif268 is also known as Early growth response protein1 (Egr1), and it is a three zinc finger of the C_2H_2 type, with a pair of cysteine residues on the β strand and a pair of histidines on theα helix, held together by a zinc ion.

The strategy to design and assemble a zinc finger protein employs either combinatorial selection of ZFPs or modular assembly of zinc finger arrays. The later approach was utilized for the purposes of this project.

The combinatorial strategy approach uses phage display to select a protein that binds specific DNA regions being studied. M13 filamentous bacteriophage is used to express on its surface pools of zinc finger proteins in which key residues of α helix involved in DNA recognition are randomised. The bacteriophage that binds to the DNA sequence of interest is then selected, amplified and sequenced to establish the identity of the amino acid residues involved in DNA recognition. Work on zinc finger phage display libraries has revealed that zinc finger DNA interactions predominantly follow a set of rules with specific amino acids targeting specific bases in a triplet configuration.

The rational design of zinc finger proteins involves the generation of modular units where each module of the three ZFP recognizes three base pairs of DNA 5'-GNN-3', 5'-ANN-3', 5'-CNN-3' and 5'-TNN-3' subunits [261-264]. The specific ZF modules can be selected using databases such as the Zinc Finger Consortium of published recognition triplet sequences. One drawback of this approach is that modules targeting certain triplets such as GTG and GCG, show low sequence specificity and can result in poor gene targeting of the assembled arrays [265].

ZF proteins can be made more specific by module - linking and recombination to increase the array size e.g. create four or six ZFPs, capable of targeting more extended 12bp and 18bp DNA contiguous sequences (Figure 1-8). For *in vivo* targeting studies this kind of set up should provide the highest level of specificity given that the probability of 18bp sequence occurring more than once in the genome is once in $68x10^9$ bp of sequence. These kind of arrangements have been observed in nature with five finger human glioblastoma protein and nine finger TFIIIA [266, 267]. Modules are linked using the highly conserved linker sequence, TGEKP.

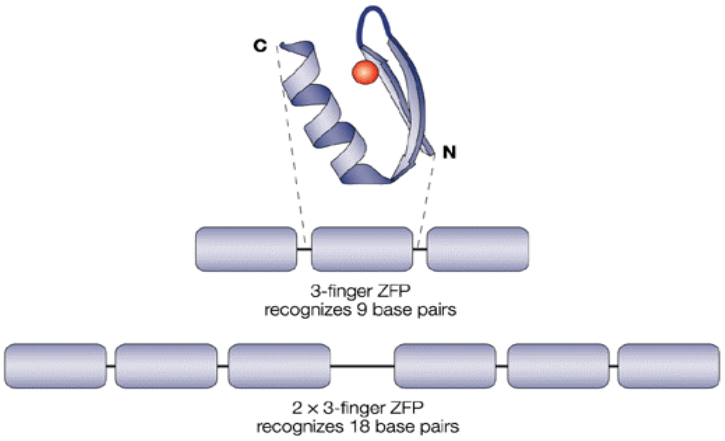


Figure 1- 8. Zinc finger assembly.

Three zinc finger modules capable of recognizing 9 base pairs can be linked *via* TGKEP linker (horizontal bar) to create six finger proteins. Figure adapted from [257].

Zinc finger proteins can be fused to a variety of different molecules such as gene repressor and activators as well as functional domains such as methylases, integrases and nucleases, enabling diverse applications [268]. Six finger polydactyl ZFproteins fused with the VP16, transcriptional activation domain and the Krupel associated box A repression domain were shown to bind with very high affinities to their target region [269].

1.4.2 Zinc fingers methyltransferase fusion

The ability to methylate endogenous gene promoter regions would facilitate further investigations into the role methylation plays in transcription/translation mechanisms associated with that gene, and enable us to mimic the early onset aberrant methylation that has been suggested to occur as part of the disease process. For the purpose of this project, *de novo* methylation will be targeted at the *CDKN2B* promoter using zinc fingers fused to methyltransferase catalytic domains with different specificities. Targeting mitochondrial DNA

using synthetic zinc finger fused to the catalytic domain of DNMT3A showed that these systems can bind selectively and cause site - specific methylation at the target sites [270]. Intracellular localization studies in this case showed mitochondrial localization of these ZFP – methylase proteins [271]. Targeted methylation has been carried out initially using non-genomic targets. Targeting constructs harbouring the TK (Human herpes virus 1 thymidylate kinase), c-Has-ras promoter and the immediate early IE175k promoter of HSV with promoter - specific zinc fingers and a catalytic domain of mouse DNMT3A and DNMT3B methyltransferases was shown to lead to gene repression. Furthermore, it was demonstrated that targeted methylation led to a decidedly distinct pattern of dense methylation observed on both strands of DNA [272].

These observations were followed by our own studies of targeting DNA cytosine methylation to a TK promoter driving a CAT reporter gene,that had been integrated into the genome of NIH3T3 cells. The TK promoter contained zinc finger recognition sites and was flanked by HpaII and HhaI sites. Methylation was targeted using four - zinc finger HpaII fusions linked by a $(Gly4Ser)_3$ linker and led to gene silencing and the formation of repressive chromatin. It was also demonstrated that this gene repression was epigenetically maintained, through many cell divisions, long after the loss of the targeted methyltransferase enzyme from the cell, presumably through the action of the endogenous cellular methylation machinery. This important study by our group showed that DNA methylation could be targeted to a specific site in the genome and proved that a repressive histone code could be initiated as a consequence [273].

In the Chapters to follow in this thesis, I will give a detailed description of the ZF specific sequences targeting the *CDKN2B* promoter region and the rationale for the selection of the target sites. Further, the cloning and assembly strategy employed to construct site-biased ZFP-methyltransferase fusions will be described. Different approaches to characterize and evaluate the specificity, binding affinity and activity of the target fusion proteins will be presented. Finally, I will demonstrate successful regional methylation delivery to the key regulatory *CDKN2B* promoter region by some of the targeted methyltransferases, which represents a significant breakthrough in the field of targeted methylation.

Chapter 2

Materials and Methods

2.1 General consumables and reagents

Phosphate Buffered Saline (PBS) Dimethyl sulfoxide (DMSO) Bacto tryptone Bacto Yeast extract Agarose Agar Amplicillin Cell culture media Glycerol Glutathione- Agarose Isopropyl β-D thiogalactoside (IPTG) X-Gal Xylenol orange Coomassie Blue R-250 N,N,N',N'-Tetramethylethylene diamine (TEMED) Glycogen Whatman 3MM paper	Sigma
40% (w/v) Acrylamide/Bis Solution 37.5:1	Bio Rad
Hybond-c extra membrane Electrogenerated chemiluninesence (ECL) detection kit Glutathione sepharose 4B	GE Healthcare
Topo-TA vector TOP10 Competent cells Trizol Novex Sharp protein marker	Invitrogen
DNA Miniprep kit RNAeasy extraction kit Genomic DNA kit Epitec Bisulfite kit HotStar taq polymerase DNAaseI RNaseA	Qiagen
Microcentrifuge 'Eppendorf' tubes 1.5ml PCR tubes 0.2ml Pipette tips Neubauer Improved Haemocytometer	Starlabs
Spin-X Centrifuge Tube Filter	Costar
DAPI 4',6, diamidino-2-phenylindole	Sigma

2.2 Restriction enzymes and polymerases

BamHI *DpnI* *DpnII* *BsmlI* *StyI* *Tsp5091* *HpaII* *MspI* *BstUI* *EcoRI* *HindIII* *NdeI* *XhoI* *BglII* *NotI* NEB buffers 1,2, 3 and 4 *EcoRI* buffer Calf Intestinal Alkaline Phosphatase (CIP) T4 Ligase Standard Ligase Buffer NEB Express Competent *E.coli* Phusion High-Fidelity DNA Polymerase Deoxyribonucleotide triphosphate (dNTP) S- adenosylmethionine (SAM)	New England Biolabs
18s VIC labelled RT-PCR RT Taqman Kit	Applied Biosystems

2.3 Tissue culture reagents and consumables

Dulbecco's Minimal Essential Medium (DMEM) Rosswell Park Memorial Institute medium (RPMI-1640) Foetal Bovine Serum (FBS) L-Glutamine Penicillin/Streptomycin Trypsin	Sigma
Tissue Culture plates and flasks Falcon tubes (15ml and 50ml) Pipettes (5, 10, and 15ml)	VWR international ltd
RT Mix for RT-PCR IcaFectin™ 441 DNA Transfection Reagent	Eurogentec
Cell Line Nucleofector Kit V and R	Amaxa
JetPEI™ Transfection Reagent	Autogen Bioclear

2.4 Solutions and Buffers

Annealing Buffer	0.1M Tris-HCl, 100mM $MgCl_2$
GST resuspension buffer (RP1)	50mM Tris-HCl (2.5ml 1M Tris-HCl pH 7.5), 1M NaCl (10ml 5M NaCl), 1mM EDTA (100µl 0.5M EDTA), 1mM EGTA (2ml 25mM EGTA), 100µM $ZnCl_2$ (5µl 1M $ZnCl_2$), 0.1mM DTT (5µl 1M DTT) and distilled water (dH_2O) to 50ml.
GST Elution Buffer (RP2)	10mM reduced glutathione: 0.0154g Glutathione, 4ml 50mM Tris-HCl, (pH 7.5) and 1ml glycerol.
Luria-Bertoni (LB) Broth	1% (w/v) bactotryptone 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl. Autoclaved and ampicillin added to a concentration of 100 µg/ml.
LB-Agar	LB broth, 1.5% (w/v) agar. Ampicillin to a final concentration of 100µg/ml
10x TBE	0.9 M Tris-HCl, (pH 8.0), 0.9 M Boric Acid, 12.5mM EDTA
10xTAE Buffer	10mM Tris-HCl (pH 8.0), 1mM EDTA
Ampicillin (1000x stock)	100mg/ml in dH_2O
Sodium Dodecyl Sulphate (SDS) loading buffer	4% SDS (w/v), 0.1 M Tris-HCl (pH 8.9), 2mM EDTA, 0.1% (w/v) bromophenol blue
RNAase A	10 mg/ml in dH_2O
STE(TEN)	0.1M NaCl,10mM Tris-HCl (pH 8.0), 1mM EDTA (pH 8.0)
MPI (miniprep buffer 1)	50mM Glucose, 25mM Tris-HCl (pH8.0), 10mM EDTA, dH_2O to 100ml. Autoclaved and stored at 4˚C.
MPII (miniprep buffer 2)	0.2M NaOH, 1% SDS (w/v)
MPIII (miniprep buffer 3)	5M KoAc 60ml, Acetic acid 11.5ml, dH_2O 28.5ml. Autoclaved and stored at 4˚C.
4x Upper buffer(stacking gel)	Tris-HCl 0.5M, dH_2O to 1L and pH to 6.8, SDS 0.4% (w/v).
4x Lower buffer (running buffer)	Tris-HCl 1.5M, dH_2O to 1L and pH to 8.8, SDS 0.4% (w/v).
Coomassie Blue Stain	230ml Methanol (MeOH), 230ml water, 40ml Acetic acid, 2g Coomassie - R 250.
Coomassie destain solution	30% MeOH, 10% Acetic acid (v/v), dH_2O 1L.
DNA loading Buffer	30% (v/v) Glycerol, 0.25% (w/v) Orange R.
DNA Elution Buffer (EB)	10mM Tris-HCl (pH 8.5)
T4 DNA Ligase Buffer	50mM Tris-HCl (pH 7.5), 10mM $MgCl_2$, 10mM dithiothreitol, 1mM ATP.
NEB Buffer 1	10mM Bis Tris Propane-HCl, 10mM $MgCl_2$, 1mM dithiothreitol (pH 7.0)
NEB Buffer 2	50mM NaCl 50mM Tris-HCl, 10mM $MgCl_2$, 1mM dithiothreitol (pH 7.9)
NEB Buffer 3	100mM NaCl 50mM Tris-HCl, 10mM $MgCl_2$, 1mM dithiothreitol (pH 7.9)
NEB Buffer 4	50mM potassium acetate 10mM Tris-acetate, 10mM magnesium acetate, 1mM dithiothreitol (pH 7.9)
BSA Storage Buffer	20mM KPO_4 50mM NaCl, 0.1mM EDTA, 5% glycerol (v/v) (pH 7.9)
EcoRI Buffer	50mM NaCl, 100mM Tris-HCl, 10mM $MgCl_2$, 0.025% Triton X -100 (v/v) (pH 7.5)
ChIP Sonication Buffer	50mM Hepes pH 7.9, 140 mM NaCl ,1 mM EDTA , 1% Triton X-100, 0.1% Na-deoxycholate , 0.1% SDS, 0.5 mM PMSF, Protease inhibitor cocktail (Roche)
ChIP Wash Buffer A	50mM Hepes pH 7.9, 500 mM NaCl ,1 mM EDTA , 1% Triton X-100, 0.1% Na-deoxycholate , 0.1% SDS, 0.5 mM PMSF, Protease inhibitor cocktail (Roche)

ChIP Wash Buffer B	20mM Tris, pH 8.0, 1mM EDTA, 250mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate, 0.5 mM PMSF, Protease inhibitor cocktail (Roche)
ChIP Elution Buffer	50mM Tris, pH 8.0, 1mM EDTA, 1% SDS,50mM $NaHCO_3$
2x PHEM Buffer	100mM PIPES, 10mM HEPS, 10mM EGTA, 2mM MgCl2 pH7.
PFA Buffer	PFA 3.75%, 10mM KOH, 5m dH_20, 5ml 2x PHEM Buffer.

2.5 Primers, probes and oligonucleotide sequences

Al the primers and oligonucleotides were purchased from either MWG Biotech (Eurofins Operon) or Sigma.

Name	Sequence	Comments
ZFP151AF1	5'CTGGCCTGTGTGGATGCGGATATGTCGGACCAAGTTGGATCGTTGAGAAAAGCGGCGATCGCAGGACTCGACAG-3'	Bottom strand
ZFP151AF2	5'TTCTCGCCTGTGTGGGTGCGGATGTGTCTGGTGAGATCCCCTGACTGACTGAAGTTACGCATGCAGA-3'	Bottom strand
ZFP151AF3	5'AAGCTTGCCGGATCCCTTTTCCCTGTATGTCGCGTCAGATCGCTTCTGTCGGCAAACTTCCTCCCACAAATG-3'	Bottom strand
ZFPB1	5'CATATGAGATCTGGAGGCGGTGAACGCCCCTATGCTTGCCCTGTCGAGTCCTGCGATCGCCGC-3'	Top strand
ZFPB2	5'CATATCCGCATCCACACAGGCCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGT-3'	Top strand
ZFPB3	ATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGGAAGTTTGCC-3'	Top strand
ZFP151BF1	5'CTGGCCTGTGTGGATGCGGATATGCCGGGCTAAGTGGTCGGATCGAGAAAAGCGGCGATCGCAGGACTCGACAG-3'	Bottom strand
ZFP151BF2	5'TTCTCGCCTGTGTGGGTGCGGATGTGCCTTGTTAAATTGGATCGGTCACTGAAGTTACGCATGCAGA-3'	Bottom strand
ZFP151BF3	5'AAGCTGCCGGATCCCTTTTCCCCTGTATGTCGGGTGAGGTTGCTTCTATCGGCAAACTTCCTCCCACAAATG-3'	Bottom strand
Zp15-all F	5'-CATATGAGATCTGGAGGCGGTGAACGC-3'	Forward primer ZFP amplification
Zp15-all R	5'-AAGCTTGCCGGATCCCTTTTCCC-3'	Reverse primer ZFP amplification
RTP15F	5'-GGACTAGTGGAGAAGGTGCG-3'	Forward RT-qPCR primer for *CDKN2B* amplification.
RTP15R	5'-TGAGAGTGGCAGGGTCT-3'	Reverse RT-qPCR primer for *CDKN2B* amplification.
RTP15PR	5'-(FAM)-CGCGGATCCCAACGGAGTCAACC-(TAMRA)-3'	RT-qPCR double dye probe for *CDKN2B* amplification.
QCZFPAF	5'-GGTCCGACATATCCGCATCCACACAGGCCAGAGCGC-3'	Forward quick change mutagenesis primer ZFPA

QCZFPAR	5’-TCTGGCCTGTGTGGATGCGGATATGTCGGACCGCGC-3’	Reverse quick change mutagenesis primer ZFPA
QCZFPBF	5’-GATCCAATTTAACAAGGCACATCCGCACCCGCGC-3’	Forward quick change mutagenesis primer ZFPB
QCZFPBR	5’-GGGTGCGGATGTGCCTTGTTAAATTGGATCCGCG-3’	Reverse quick change mutagenesis primer ZFPB
P15LF1	5’-GCGCAAGCTTATGACAGGTTCAGCCTGTTTAAT-3’	Forward primer for *CDKN2B* amplification.
P15LR1	5’-GGCCGTAAACTTAACGACACTCCTCGAGGCGC-3’	Reverse primer for *CDKN2B* amplification.
P15SF1	5’-GCGCAAGCTTAGCCACCAACGTCTCCACAGTGAA-3’	Forward primer for *CDKN2B* amplification (734bp)
P15SR1	5’-TGGGCATGCCCTTGTTCTCCTCGCTCGAGGCGC-3’	Reverse primer for *CDKN2B* amplification (1206bp)
1a1a-4z-limus-5 (EMSA)	CTAGAGGTGGCTTCTGCGTCGCTTAGGCATATGCCGGCGCAGA GTACTGCCGCAGAAGGACGACGGGTTACA	First ZFP oligo strand.
1a1a-4z-limus-3 (EMSA)	GATCTGTAACCCGTCGTCCTTCTGCGGCAGTACTCTGCGCCGG CATATGCCTAAGCGACGCAGAAGCCACCT	Second ZFP oligo strand.
6-litmus-3 (EMSA)	GATCTGTAAGCCCACTCCCCACAGCCACTGCGCCGGCATATGC CTAAGCGCGTGGGCGAGCGTGGGCGACCT	First NS oligo strand.
6-litmus-5 (EMSA)	CTAGAGGTCGCCCACGCTCGCCCACGCGCTTAGGCATATGCCG GCGCAGTGGCTGTGGGGAGTGGCGTTACA	Second NS oligo strand.
3'HPA 1 BIS F3	GGA GGG TAA TGA AGT TGA GTT TAG G	COBRA 3’ 1st round
3' HPA 1 BIS R2	CTC CTT TTC TCC CCA ATT CAA TC	COBRA 3’ 1st round
3' HPA 2 BIS F1	GTT GAG TTT AGG TTT TTT AGG AAG G	COBRA 3’ 2nd round
3' HPA 2BIS R5	TCT CCC CAA TTC AAT CTA TTC C	COBRA 3’ 2nd Round
F1	AGT TTA AGG GGG TGG GGA GA	COBRA 5’ round
R1	CCA CCC TAA CCA AAC CCT CAT C	COBRA 5’ round
F2	GAG TTT AAG GGG GTG GGG AGA	COBRA 5’ round
R7	AAA CTA TCC CAC CTT CTC CAC TAA TCC	COBRA 5’ round
TET1 F	CAG GAC CAA GTG TTG CTG CTG T	TET PCR

TET1 R	GAC ACC CAT GAG AGC TTT TCC C	TET PCR
TET2 R	AGA GAA GGA GGC ACC ACA GGT T	TET PCR
TET2 F	GCT TAC CGA GAC GCT GAG GAA A	TET PCR
TET3 F	CCA CAA GGA CCA GCA TAA CCT C	TET PCR
TET 3R	CTC GCT ACC AAA CTC ATC CGT G	TET PCR
CHIP CTCF F	ACA GGT TCA GCC TGT TTA AT	ChIP
CHIP CTCFR	ACA GTA AAT CTT TGT TTG CAC TC	ChIP
GAPDH Rev 5'	TCG AAC AGG AGG AGC AGA GAG CGA	ChIP
GAPDH For 5'	TAC TAG CGG TTT TAC GGG CG	ChIP
ET1 REV	GCA GAG GCG CGT CTG ACT T	ChIP
ET1 FW	TGC ACG TTG CCT GTT GGT	ChIP

2.6 Cloning Plasmids

Expression plasmid	Plasmid description	Vector host
pGEX5mcs	GST purification and expression plasmid	Bacterial cells
pBL-CAT	Expression plasmid with CAT reporter gene	Mammalian cells
pCMV	Expression plasmid with CMV promoter	Mammalian cells
Litmus 28i vector	General cloning plasmid	Bacterial cells

2.7 Antibodies

Antibody	Antibody description	Use
FLAG®-TAG Monoclonal antibody	Sigma F-3165 .	**WB/IF**
Rat anti mouse IgG1	BD Pharmingen™ 553443	**IF**
Alexa Fluor® 633 Goat anti-rabbit IgG	Life Technologies A21071	**IF**
Rabbit polyclonal anti VEZF1 antibody	Santa Cruz Biotechnologies sc-98278X	**ChIP**
Anti-GAPDH goat polyclonal antibody	Chemicon MAB374	**WB**
Polyclonal rabbit anti goat IgG -HRP	Dako PO160	**WB**
Polyclonal Rabbit 5-hmeC	Active Motif 39769	**IF**
Alexa Fluor® 488 – rabbit polyclonal antibody	Life Technologies A11008	**IF**
Alexa Fluor® 633 rabbit polyclonal antibody	Life Technologies A21070	**IF**
Alexa Fluor® 488rabbit anti-mouse IgG (H+L)	Life Technologies A11059	**IF**
Anti- methyl cytosine mouse monoclonal antibody	Eurogentec BI-MECY-0100	**IF**
Polyclonal Goat anti mouse IgG – HRP	Dako PO447	**WB**
Polyclonal Goat anti rabbit IgG – HRP	Dako EO432	**WB**

2.8 DNA and RNA procedures

2.8.1 DNA electrophoresis

Agarose powder was dissolved in 1x TAE, or 1x TBE buffer to 1-2% (w/v) by boiling in a microwave oven. After cooling to ~ 50°C, ethidium bromide was added to a final concentration of 0.5µg/ml and the gel solution was then poured into a tray with removable combs and left to set at room temperature. Loading dye was added to DNA samples prior to gel loading and running, typically at 80V for 20-50 minutes (depending on the size range of DNA to be analysed size). DNA was visualized on a UV transilluminator. DNA ladders of known sizes were used to determine the size and the approximate concentration of the DNA.

2.8.2 DNA Purification

Agarose gel slices containing the desired DNA fragment were placed in a Spin-X column, 150µl dH_2O was added, the tube frozen at –20˚C for 30 minutes. The sample was then centrifuged at 14,000xg in a microcentrifuge (which was the speed for all subsequent centrifugation steps) for 10 minutes. 300µl of phenol/$CHCl_3$/IAA (25:24:1) was added to the filtrate and the solution vigorously vortexed then centrifuged for 2 minutes. The upper phase layer was carefully removed to a fresh tube for a further phenol extraction. DNA was precipitated by addition of 1µl of glycogen, briefly mixing and then adding three volumes of 100% ethanol (EtOH) (-20°C). After storing at -20°C for 30 minutes, the solution was centrifuged for 20 minutes. EtOH was carefully removed, the DNA pellet washed with 0.5ml of 70% v/v EtOH and centrifuged for 5 minutes. The EtOH was removed by aspiration, the pellet air dried and resuspended in 50µl EB buffer. The purified DNA was subsequently stored at - 20°C.

2.8.3 Genomic DNA extraction and Bisulphite modifications

The Genomic DNA Kit (Qiagen) was used for genomic DNA extractions. The EpiTec ® Bisulfite Kit (Qiagen) was used for bisulfite modification. Manufacturer's instructions were followed for both procedures. Briefly, genomic DNA was extracted from several human cell lines and the DNA concentration determined at 260nm using a Nanodrop 1000 (Thermo Scientific). Genomic DNA (1µg) was converted with sodium bisulphite treatment using EpiTec® bisulphite kit (Qiagen) in a 140µl total volume reaction containing 85µl of Bisulphite Mix, 35µl of DNA Protect Buffer and DNA. The bisulphite conversion was performed in a thermocycler with following settings:

Step	Time	Temperature (°C)
Denaturation	5 minutes	99
Incubation	25 minutes	60
Denaturation	5 minutes	99
Incubation	3 hours	60
Denaturation	5 minutes	99
Incubation	6 hours	60
Hold	Indefinite	20

Following bisulphite conversion, the PCR tubes were centrifuged briefly and the sample transferred to a new microcentrifuge tube. Buffer BL (560µl) containing 10µg/ml carrier RNA was added to the samples, mixed by vortexing, centrifuged briefly and transferred to an EpiTec spin column. The column was spun at 13,000xg (which was the speed used for all following centrifugation) and the flow through discarded. The column was washed with 500µl of BW Buffer. DNA desulfonation was carried out by addition of 500µl of Buffer BD and incubation for 1 hour at 37°C. The column was spun again to remove the flow through and washed twice with 500µl of the BW Buffer. DNA was eluted in 20µl of dH_2O.

2.8.4 Combined Bisulphite Restriction Analysis (COBRA)

Bisulphite converted DNA was PCR amplified using a nested PCR approach which involved using two sets of primers in two successive PCR reactions. In the first reaction typically multiple PCR products are generated, visualised as a smear on an agarose gel. This is used as a template for a second PCR, using primers which are "nested" within the first set, which typically generates a single product. The typical cycling conditions are given in the table below. This PCR product was digested with methylation sensitive enzymes (Combined Bisulfite Restriction Analysis (COBRA) in order to determine approximate methylation status. Specifics for different target products are explained in more detail in Chapter 4 of this thesis. Additionally bisulphite PCR libraries were directly sequenced (commercially) or TOPO®-TA cloned if evaluating individual clones.

Cycles	Temperature	Time	Requirement
1	95°C	5 minutes	Enzyme activation
35	95°C	30 seconds	Denaturation
	54/57°C	2 minutes 40 seconds (1st round PCR), 50 seconds (2nd round PCR)	Primer annealing temperature
	72°C	1 minute 20 seconds	Elongation
1	72°C	10 minutes	Time to fill in any nucleotides missing from ends

2.8.5 Restriction Digests

Typically restriction digests contained 2µl of 10x restriction enzyme buffer, 2µl 10x Bovine Serum Albumin (BSA, 1mg/ml), 0.5-1µl of restriction enzyme(s) (5-10 units/µl), 1µl RNAase (1mg/ml) and 3µl of DNA (approximately 1µg) in a total volume of 20µl. All digestions were incubated between 1 and 2 hours at the appropriate temperature. Vector DNA to be used for cloning was additionally treated with 1µl (10 units/µl) of Calf Intestinal Phosphatase (CIP) during the restriction process. CIP treatment of the vector DNA removes 5' phosphate groups, thereby reducing self-ligation.

The table below lists single and double digests used for different cloning procedures.

Enzyme(s)	**Buffer**	**Cloning procedure**
EcoRI	EcoRI buffer	TOPO®-TA
HindIII/NdeI	NEB buffer 2	pGEX/ZFP subcloning
BamHI/NotI	NEB buffer 3	pGEX/ZFP/HpaII subcloning
BamHI/HindIII *BglII/HindIII*	NEB buffer 2	pGEX/ZFP A+B 6 ZFP subcloning
XhoI/HindIII	NEB buffer 2	pBLCAT/*CDKN2B* subcloning
XbaI/BglII	NEB buffer 2	Litmus 28i vector cloning
NdeI/AccIII	NEB buffer 2	pCMV-SssI ZFP subcloning

2.8.6 DNA cloning

2.8.6.1 pGEX ZFP Cloning

Assembled zinc fingers (see 2.8) were excised from TOPO®-TA vectors using *R.NdeI/Hind III* double digests. The pGEX -5mcs plasmid was digested using the same set of enzymes. DNA was separated by agarose gel electrophoresis and the fragments of interest were gel extracted and ligated as described in section 2.7.2. This generated the three zinc finger vectors ZFPA and ZFPB.

The six zinc finger vector was assembled as a fusion of ZFPA and ZFPB modules. The pGEX ZFPA vector was *R. BamHI/HindIII* restricted to accept the *R. BglII/HindIII* restricted ZFPB fragment. For M.HpaII cloning for example, the pGEX ZFA plasmids were double digested using

R. Bam HI/NotI enzymes. This was used as a platform for subsequent swapping of ZF components.

2.8.6.2 pBLCAT *CDKN2B* Cloning

The pBLCAT plasmid was digested with *R. XhoI/HindIII* enzymes and the *CDKN2B*-730 *R. XhoI/HindIII* fragment was subcloned into this vector.

2.8.6.3 TOPO®-TA Cloning

All TOPO vector cloning procedures were carried out using the TOPO®-TA cloning kit and TOP10 competent cells. The typical reaction volume was 6µl, consisting of 1µl Salt solution (1.2M NaCl, 0.06M $MgCl_2$), 1µl (5ng) TOPO- TA vector and 4µl DNA fragment (~100ng). The reaction was incubated for 5 minutes at room temperature then added to 50µl of competent TOP10 cells for transformation, as described in section 2.7.8.

2.8.6.4 Litmus 28i Cloning

The Litmus 28i vector was digested using *R. XbaI/BglII* enzymes. Oligonucleotides harbouring either zinc finger recognition sites or non-relevant specific DNA sequences were heated for 5 minutes at 100˚C in annealing buffer and allowed to anneal by cooling to room temperature overnight. Annealed oligonucleotides with XbaI/BglII overhangs were subcloned into the Litmus 28i vector. Positives clones were confirmed by sequencing.

2.8.7 DNA Ligations

DNA ligations were typically carried out in a 20µl reaction volume containing 2µl Ligase buffer, 1µl T4 Ligase (10 units/µl) and a ratio of vector to insert DNA typically between 1:3 to 1:5. Ligation negative controls containing no insert DNA were also set up. The reaction was incubated overnight at 4°C.

2.8.8 DNA Transformations

Chemically competent TOP10 cells (genotype F- *mcr*A Δ (*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ (*araleu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G) were thawed on ice. 50µl of the competent cells were added to the ligation mix. Following incubation on ice for 30 minutes, the reaction was subjected to heat-shock for 30 seconds at 37^{o}C, followed by an additional 2 minutes on ice. Transformed cells were supplemented with 250µl of LB media and incubated for a further 1 hour at 37^{o}C with shaking. The cells were then plated out on LB-Amp agar pre-warmed plates and incubated at 37^{o}C overnight. For blue / white selection of colonies 80µl of 40mg/ml X-gal was plated out 10 minutes prior to seeding. The TOPO-TA vector is unable to re-circularize due to its 3'dT overhangs. White colonies result from DNA being inserted into the vector thus disrupting the *β-galactosidase* gene. Blue negative colonies are a result of X-gal staining and indicative of a functional *β-galactosidase* gene.

2.8.9 Small and large scale DNA preparations

Selected individual colonies were inoculated in 5ml of LB media overnight at 37°C. The cells were pelleted by centrifugation at 2000xg for 3 minutes and the pellet re-suspended in 100µl of MP I solution. The cells were then lysed in 200µl of MP II lysis solution and incubated for 5 minutes at room temperature. The mixture was neutralised by the addition of 150µl of MPIII solution and the lysate was then clarified by centrifugation for 10 minutes at 13,000xg. The supernatant was removed, mixed with 400µl of phenol/$CHCl_3$/IAA, vortexed and centrifuged for 3 minutes at 13,000xg. The separated aqueous upper phase layer was transferred into a fresh 1.5ml tube and ethanol precipitated. The pellet was washed in 70% (v/v) EtOH, allowed to air dry and re-suspended in 40µl of dH_2O.

For large scale plasmid DNA preparations (used for all transfections), a Qiagen plasmid Maxi Kit or GenElute™ HP Plasmid Kit (Sigma Aldrich) were used according to manufacturer instructions. Selected individual colonies were inoculated in 150ml of LB media with ampicillin, overnight at 37°C in a shaking incubator at 150rpm. The culture was centrifuged for 15 minutes

at 4000xg in a chilled centrifuge and the pellet re-suspended in re-suspension buffer, lysed with lysis Buffer, neutralised and extracted according to the manufacturer instructions.

2.8.10 RNA extraction and cDNA synthesis procedure

Cell lines (typically $1x10^6$ cells) were harvested and pelleted. The cells were then homogenized with 750µl of Trizol and incubated at room temperature for 5 minutes. The lysed sample was mixed with 200µl of chloroform and centrifuged at 12,000xg for 15 minutes. An equal volume of 70% (v/v) ethanol was added to the separated upper aqueous phase in a fresh tube and applied to an RNA binding column RNAeasy Kit (Qiagen). The column was centrifuged for 15s at 8000xg and the eluate discarded. DNAaseI (80µl) was added to the column followed by a 15 minute room temperature incubation. Subsequent washing with 700µl of Qiagen wash buffer was followed by elution of purified RNA with 20µl of RNAase free water, by centrifugation for 1 minute at 15000xg.

Extracted RNA was reverse transcribed using the RT –Taqman™ Applied Biosystems Kit.

Reagent	**Volume**
10mM dNTP Mix (2.5mM each)	2.1µl
Random hexamer (50µM)	0.56µl
RNA	2.0µl (1-5µg)
25mM $MgCl_2$	2.31µl
10x RT buffer	1.05µl
RNAase Inhibitor	0.2µl
MultiScribe ™Reverse Transcriptase (50 units/µl)	0.66µl
dH_2O	1.55µl
Total volume	10µl

The reverse transcription reaction was incubated for 10 minutes at 25°C, followed by 30 minutes at 48°C and 5 minutes at 95°C.

2. 9 Zinc Finger Assembly -Splicing by Overlap Extension PCR

Splicing by overlap extension (SOE) was used for zinc finger assembly. DNA coding for zinc finger proteins specific for differing regions of the *CDKN2B* promoter were designed using published zinc finger recognition code information software [274], based on extensive published zinc finger : DNA recognition data. Briefly, six oligonucleotides coding for the alpha helix and beta sheet regions of zinc fingers were designed to contain 15 base complementary overlap regions. Oligonucleotides were annealed, extended, and the flanking strands were 'filled in' using four cycles of RT PCR, followed by 25 cycles of primer specific PCR.

PCR reactions were set up according to the table below.

Reagent	**Volume**
10mM dNTP Mix	1µl
Phusion Polymerase (2 units/µl)	1µl
Oligonucleotide	100ng/µl
10x HF® buffer	5µl
dH_2O	To make up a total volume of 50µl

PCR cycling conditions used were:

Cycles	**Temperature**	**Time**	**Requirement**
5	95°C	5 seconds	Enzyme activation
	35°C	10 seconds	RT step
	72°C	20 seconds	Elongation
Reaction was stopped and 1µl of reverse and forward primers added. The reaction was re-started using the following conditions.			
35	95°C	5 seconds	Denaturation
	55°C	10 seconds	Primer annealing temperature
	72°C	20 seconds	Elongation
1	72°C	10 minutes	Time to fill in any nucleotides missing from ends

The PCR product was gel purified and prepared for TOPO®-TA cloning as follows: 5µl of PCR product, 1µl Taq polymerase (5 units/µl), 2µl Taq Standard Buffer, 1µl dNTP at 10mM, and 11µl dH_2O, was incubated at 72°C for 20 minutes and TOPO cloned for sequence verification.

2.9.1 Quick-change mutagenesis

Primers coding for the region of interest were gel purified and set up in the following reaction:

Reagent	**Volume**
10mM dNTP Mix	1µl
Phusion Polymerase (2 units/µl)	1µl
DNA	30µl (100ng/µl)
10X HF® buffer	5µl
Forward primer 10µM	1µl
Reverse primer 10µM	1µl
dH_2O	To make up a total volume of 50µl

PCR programme used was:

Cycles	**Temperature**	**Time**	**Reaction**
18	95°C	5 seconds	Enzyme activation
	55°C	10 seconds	RT step
	72°C	3 minutes and 20 seconds	Elongation

The product was then incubated with 6µl of Buffer 4, 1µl 100x BSA and 1µl DpnI enzyme for 40 minutes at 37°C, gel purified, and then used to transform *E.Coli*. More details on the quick-change procedure are provided in Chapter 3, Section 3.2.1. The same chapter also contains a detailed procedure for Zinc finger expression and purification Section 3.1.3.

2.10 Cell culture

The following cell lines were used in this study:

Cell Line	Origin	Culture conditions
HL60	Human promyelocytic leukaemia	Suspension
Jurkat	Human leukemic T cell lymphoblast	Suspension
Raji	Human Burkitt's lymphoma	Suspension
HeLa	Human cervix epitheloid carcinoma	Adherent
A549	Human Caucasian lung carcinoma	Adherent
293(HEK)	Human Embryonic Kidney 293 cells	Adherent

2.10.1 Suspension cell culture

Frozen stocks of non-adherent cell lines were revived in RPMI 1640 media supplemented with 20% (v/v) Foetal Bovine Serum (FBS), L-glutamine (2mM final) penicillin (2,000 units/ml). Suspension cells were maintained at 0.2-0.5x10^6 cells/ml in 75cm^2 flasks. The FBS concentration was gradually reduced to 10% (v/v), when the cells became confluent and viable as determined by Trypan blue count. Cells were harvested by centrifugation at 200xg for 10 minutes.

2.10.2 Adherent cell culture

Adherent cell lines were initially revived in 25cm^2 tissue culture flasks and further propagated in 75cm^2 tissue culture flasks. The cells were routinely split at 1:3 dilution when they had reached ~80% confluence. Cells were washed with 10ml PBS following media removal, trypsinized with 0.6 ml of EDTA-trypsin and incubated at 37^{o}C to allow detachment. Trypsinization was stopped with 10ml of DMEM media supplemented with 10% (v/v) FBS, L-glutamine (2mM final) penicillin (2,000 units/ml) and the cells re-seeded at the appropriate ratio.

The same procedure was applied for the routine passage and harvesting of transfected cell samples. These well were grown in either 24 well, 12 well, 6 well plates or 10 cm dishes depending on the assay time point. Volumes of PBS and Trypsin were adjusted accordingly.

Harvesting of disaggregated adherent cells involved centrifugation at 200xg for 10 minutes to pellet the cells. The supernatant was removed, the pellet washed in 1ml of 1x PBS and pelleted again by centrifugation at 200xg for 10 minutes. The pellets were frozen at -20°C until required for analysis.

2.10.3 Cell Transfections

To establish the best transfection procedure for the cell lines used, several transfection methods were evaluated and the transfection efficiency determined using FACS analysis of GFP transformed cells. Transfection reagents evaluated were Nucleofector Kit V for HL60 (Amaxa), Nucleofector Kit R for HeLa (Amaxa), IcaFectin™ 441 Transfection reagent (Eurogentec) and JetPEI™ Transfection reagent (Autogen Bioclear).

2.10.3.1 Amaxa Nucleofection – Electroporation Method

The Amaxa Cell Line Nucleofector™ Kit V or R was used to transfect either plasmid constructs (maxi-prep 1-5µg) or siRNAs, using the Nucleofector® Device from Amaxa (Lonza). The transfection solution was prepared by mixing Supplement solution (0.5ml) and Nucleofector reagent (2.5ml). The solution was pre-warmed for 30 minutes at room temperature. Supplemented media (RPMI 1640 media or DMEM) was pre-warmed at 37°C. Cells were counted and adjusted to $2x10^6$ cells/ml, for each sample. The cells were then centrifuged at 200xg for 5 minutes, the supernatant carefully removed and the pellet re-suspended in 100µl of the Nucleofector mix. 1µg of pGFP (Amaxa) or expression plasmid was then added to the cell suspension, which was subsequently transferred carefully to avoid air bubbles, into the Amaxa certified cuvette. The cuvette was capped and placed into the Nucleofector Device. The correct

Nucleofection program was selected for each specific cell type, for either high efficiency or high viability. The transfected cell preparation was then supplemented with 0.5ml of pre-warmed media and the samples transferred to 4ml of media in 6-well plates and incubated at 37°C for 24 hours. Following incubation, the transfected cells were centrifuged for 5 minutes at 200xg, washed once in 1ml PBS, centrifuged again at 200xg for 5 minutes and re-suspended in 1ml PBS. The cells were then analysed for GFP transfection efficiency *via* FACS analysis.

2.10.3.2 IcaFectin™ 441/Lipofectamine™ Transfections

Cells were grown in 24 well plates to ~50% confluence ($1x10^5$ cells/ml) 24 hours prior to transfection and moved to antibiotic free medium 30 minutes prior to transfection. 20µl of IcaFectin™441 (Eurogentec) was incubated with 1µg pGFP (Amaxa) plasmid for 30 minutes at room temperature. The DNA: reagent solution was then added to the cells drop-wise and incubated for 24 hours at 37°C, 5%CO_2. The cells were then analysed for GFP transfection efficiency *via* FACS analysis.

2.10.3.3 JetPEI™ Transfections

Cells were prepared for transfection as described in the previous section. pGFP (1µg) plasmid was incubated with 50µl of JetPEI™ reagent for 10 minutes at room temperature. The DNA: reagent solution was then added to the cells drop-wise and incubated for 24 hours at 37°C. Cell preparation and analysis was carried out as described in previous section.

2.10.3.4 CaBES Transfections

The CaBES method was used to transfect adherent cell lines only. 24 hours prior to transfection, cells were grown in 10cm tissue culture plates to ~60% confluence and the media changed 1 hour prior to transfection. The transfection mixture was prepared by incubating pDNA expression vector (10µg) or pGFP (1µg) Amaxa, with 500µl of 2XBES and made up to 875µl with sterile dH_2O. The mixture was gently vortexed and centrifuged briefly.125µl of 1M $CaCl_2$ was added to the mixture slowly drop wise whilst swirling the tube at room temperature. Following 20 minutes incubation, the transfection mixture was added drop-wise to plated cells, gently swirled to ensure even distribution of the mixture, and left overnight at 37°C, 5% CO_2.

2.11 EMSA assay

Complementary oligonucleotides harbouring zinc finger recognition sites or non- specific DNA were heated for 5 minutes at 100°C in annealing buffer and allowed to cool to room temperature overnight. Annealed oligonucleotides were designed to have 5' overhangs to allow labelling via Klenow polymerase enzyme action.

2.11.1 Labelling oligonucleotides

The annealed oligonucleotides were labelled according to the procedure described in the table below.

Reagent	**Volume µl**
Oligonucleotides (100ng)	4
NEB Buffer 2	3
2mM dATP/dGTP/ dTTP (each)	1 of each
Klenow fragment 5U/µl	1
[^{32}P] dCTP (3 000 Ci/mmol)	3
10x BSA	3
d H_2O	To make total volume of 30

After incubated at 37°C for 1 hour, the labelling reaction was stopped by adding 100µl of 1x STE, 400µl of phenol/$CHCl_3$/IAA, mixing and then centrifuging for 2 minutes at 13,000xg. The upper aqueous layer was transferred into a spin column containing Sephadex G50 resin (pre-equilibrated in STE) and centrifuged for 4 minutes at 2500xg. Eluted samples were stored at 4°C for EMSA assays.

2.11.2 EMSA assay procedure

Labelled probes were initially run on a 4% DNA poly acrylamide gel (19:1, Acryl: Bis) in 1x TBE in order to determine relative labelling efficiencies between probes. Gels were pre run for 30 minutes at 80V prior to sample loading. A master mix consisting of labelled oligonucleotides was set up for 10 samples to contain: 10x BSA, 25µl; Poly dI:dC/dI:dC (100ng/ml), 25µl; 5x McNabind zinc finger buffer and 25µl probe, in a final total volume of 125µl. This mix was aliquoted into 10 separate tubes and GST fusion proteins were added to individual tubes according to the dilution table below. The binding reaction was allowed to proceed for 30 minutes at room temperature. After this time, 2µl of orange loading dye was added and samples run at 80V, until the dye front reached the bottom of the gel (approximately 40 minutes).

The gel was removed, placed on Whatman filter paper, sprayed with water, wrapped in Saran wrap, dried under vacuum for 1 hour at 80°C and exposed to a Fuji Phosphorimager plate for 1 hour and then to x-ray film at -70°C overnight.

Reaction	**Probe (add 9 µl of master mix)**	**Protein µg**	**H_2O µl**
1	Probe 1	0	11
2	Probe 1	0.008	10
3	Probe 1	0.006	8
4	Probe 1	0.004	6
5	Probe 1	0.08	10
6	Probe 1	0.06	8

7	Probe 1	0.04	6
8	Probe 1	0.8	10
9	Probe 1	0.6	8
10	Probe 1	0.4	6

2.12 The *In vitro* methylation activity assay

The *in vitro* methylation activity assay reaction mixture was set up as follows:

5x McNabind zinc finger buffer	20µl
10x BSA	10µl
SAM (32mM)	1.5µl
LI-ZFP or LI-NS	2.5µg
dH_2O	To make 100µl

ZF/HpaII protein fusion was added to separate tubes at the levels shown in the table below and 9 µl of the above reaction mixture was added to each of the tubes:

Tube	1	2	3	4	5	6	7	8	9	10
Protein(µl) (2µg/µl),	0	1	2	3	4	5	6	7	8	9
dH_2O (µl)	11	10	9	8	7	6	5	4	3	2

Following a 45 minute incubation at 37°C, the reaction was terminated by heat killing the protein at 100°C for 15 minutes. Following this incubation, 2.5µl NEB Buffer 2 and 1µl *R.HpaII* restriction

enzyme was added and the reaction incubated overnight at 37°C. The restriction digests were then separated by electrophoresis through a 2% (w/v) TBE gel at 95V for 60-80 minutes.

2.13 Construction of *CDKN2B* reporter construct

HL60 genomic DNA was used as a template for these PCR reactions. The primers were designed to amplify various *CDKN2B* promoter regions and to introduce XhoI/HindIII restriction sites, to enable subsequent subcloning into pBLCAT. The first primer pair amplified the region -253 to + 472 from the Transcription Start Site (TSS) and the second was -791 to + 398 (for more details see Chapter 3, Figure 3-2). The PCR conditions were as follows:

Reagent	**Volume**
10mM dNTP Mix	2µl
HotStarTaq™ polymerase (Qiagen)	1µl
HL60 genomic DNA	0.5µl (1µg/µl)
PCR buffer	5µl
Forward primer (100µM)	1µl
Reverse primer (100µM)	1µl
$MgCl_2$ (25mM)	3µl
dH_2O	To make up a total volume of 50µl

Cycles	**Temperature**	**Time**	**Reaction**
1	95°C	15 minutes	Enzyme activation
35	94°C	30 seconds	Denaturation
	57°C	45seconds (ramp rate 20)	Primer annealing temperature
	72°C	60 seconds	Elongation
1	72°C	10 minutes	Elongation period to fill in nucleotide ends.

PCR fragments were gel purified, restricted, re-purified and subcloned into pBLCAT as *R. XhoI/HindIII* fragments.

2.14 Gene Expression Analysis

2.14.1 *CDKN2B* RT-qPCR

Real time PCR was performed using an ABI Prism 7900HT with Quantitative PCR Master Mix - UNG with ROX. The following PCR conditions were set up:

Reagent	**Volume**
Quantitative Master Mix Plus UNG with ROX	12.5µl
F Primer (10mM)	1µl
R Primer (10mM)	1µl
dH_2O	5.5µl
cDNA	5µl
total	25µl
Real Time PCR programme:	

94°C	4min	
94°C	30sec	40 cycles
61°C	35sec	
72°C	45min	

2.14.2 $p15^{INK4B}$ Western Blot

2.14.2.1 SDS PAGE Protein Gel preparation

A 15% acrylamide gel consisting of resolving and stacking gel layers was made according to the table below. Samples and low molecular weight marker were loaded at 30µl per well and the gel run at 80V in 1x SDS running buffer, until the dye front passed the stacking gel layer (1-2 hours). The voltage was then increased to 120V until the bromophenol dye reached the bottom of the resolving gel. The gel was stained overnight with Coomassie blue stain, and destained as required.

Reagent	Running gel	Stacking gel
40% (w/v) Acrylamide (37.5:1)	3.75ml	1.125ml
Running gel buffer	2.5ml	
Stacking gel buffer		1.25ml
dH_2O	3.75ml	2.625ml
	Mix and add immediately prior to gel loading	
10% (w/v) APS	100µl	50µl
TEMED	10µl	5µl

2.14.2.2 Protein sample preparation

Frozen cell pellets (at $2x10^5$-$1x10^6$cells/ml) were resuspended in 20µl of 2xSDS loading buffer, vortexed and heated for 5 minutes at 100°C to denature the protein. Proteins were separated by SDS–PAGE.

2.14.2.3 Western blotting

Western blotting was performed following the SDS-PAGE resolution. The SDS PAGE gel was overlaid with nitrocellulose-coated nylon membrane and sandwiched between Whatman filter papers and sponge pads, pre-soaked with Tris-glycine, 20% (v/v) methanol. Proteins from the gel were transferred onto the nitrocellulose membrane by electrophoresis at 20A, 40V for 60 minutes.

2.14.2.4 Protein detection

The nitrocellulose membrane was blocked with 5% (w/v) dried skimmed milk (Marvel) diluted in PBS-0.05% (v/v) Tween 20 for 1 hour. Following a PBS-Tween-20 rinse, the membrane was incubated overnight in anti-p15 mouse monoclonal antibody (1:500), diluted in 5ml PBS 0.05% (v/v) Tween 20 and 1% (w/v) dried skimmed milk, at 4°C. The membrane was then washed three times for 5 minutes in PBS-0.05% (v/v) Tween-20. This was followed by a 1 hour incubation with anti-mouse Horse Radish Peroxidase (HRP) conjugated secondary antibody (1:2000) in 1% (w/v) dried skimmed milk diluted in PBS-0.05% (v/v) Tween. The membrane was

subsequently washed three times in PBS-0.5% (v/v) Tween-20 for 5 minutes, incubated for 5 minutes in ECL-plus detection reagent, wrapped in saran wrap and then exposed to X-ray film.

2.15 Immunofluorescence Staining

Transfected cells were trypsinized 24 hours post transfection, suspended in 400µl of complete media and plated out onto Poly-L-lysine coated slides. Cells were allowed to attach to the slides by incubation for 24 hours at 37°C. Excess media was removed *via* aspiration following incubation and the cells were fixed with 400µl of PFA for 15 minutes at RT. Slides were then washed three times with PBS and lysed in a solution of 0.5% Triton for 10 minutes. This was followed by another three washes with PBS and blocking using 10% donkey serum in PBS for 30 minutes. Serum was removed *via* aspiration and primary antibody was added to the slides and incubated O/N at 4°C. Following this incubation, antibody solution was removed and the slides were washed three times with PBS and blocked again using 10% donkey serum in PBS for 30 minutes. Serum was removed under suction and a secondary anti-conjugate antibody diluted in PBS with 10% donkey serum, was added to the slides and incubated for 2 hours at room temperature in the dark. Slides were then washed three times with PBS and stained with 400µl of DAPI solution. The slides were incubated for 10 minutes in the dark at RT, fixed with mounting solution, covered with a coverslip and left overnight. Slides were viewed using a fluorescent microscope with oil immersion.

2.16 Formaldehyde cross linking

$1x10^5$HeLa cells were seeded 24 hours prior to transfection in 6cm TC plates and allowed to reach ~70% confluence. The following day, cells were transfected using the CaBES method described previously. GFP expression was used to measure transfection efficiency, recorded at 24, 48 and 72 hours post transfection. Cells were trypsinized and transferred to a 10cm dish 48 hours post- transfection and left to incubate at 37°C for a further 2 days. At day 4 cells were trypsinized, washed with PBS two times and suspended in 37 ml of complete media with addition

of 1% (w/v) formaldehyde. The cells were then incubated for 10 min at room temperature for the cross- linking reaction to take place. The reaction was stopped by adding glycine to 45mM final concentration. The cells were then placed on ice and washed three times with 20 ml of ice-cold PBS/0.5mM PMSF/Protease inhibitor mix (Roche). Following the final wash, cells were pelleted for 5 minutes at 1000xg, flash frozen on dry ice and stored at -80°C.

2.17 Chromatin Immunoprecipitation (ChIP)

Cell pellets were suspended in 400µl of ChIP Sonication Buffer and sonicated in Bioruptor® (Diagenode). To assess chromatin status following sonication, 50µl samples of sonicated material were removed and reverse cross- linked by adding 0.25µl of Proteinase K (25 mg) and 0.5µl of RNAase Solution (10 mg/ml) and heating O/N at 65°C. DNA was subsequently purified using a PCR clean up column and visualized by agarose gel electrophoresis. Samples were only used for ChIP analysis if the DNA was in the region between 200-1000 nt, with the lower molecular weight DNA being the most dominant species. The lysate was pre-cleared by incubating with 50µl Protein A Sepharose beads and 10µg yeast t-RNA for 2 hours at 4°C with constant rotation. The samples were then centrifuged at 300xg for 5 min at 4°C and the pre-cleared soluble chromatin supernatant was removed. 50µl of the total soluble chromatin, which is 1/20th of the amount used per IP was removed and stored at -20°C, for the latter preparation of input DNA. The protein concentration of each sample was determined by BCA. In total 20µg of nucleosome material was used for each ChIP and the volume was adjusted to 1 ml with ChIP incubation buffer. The samples were then mixed with 5µg of the antibody of interest (VEZ-F1, FLAG) or 5 µg of polyclonal rabbit anti-mouse immunoglobulin (IgG) as a negative control. The ChIP complex was then rotated at 20-30 rpm for 2 hours at 4°C. Following incubation, 40µl of pre-washed magnetic protein A beads (AdemTech 04230-AD) were added and incubated for 12-16 hours at 4°C with agitation. (Previous studies in this lab have shown no requirement for pre-blocking of these beads in ChIP experiments). All subsequent washes were carried out at room temperature. The reaction tubes were placed in a magnetic rack (Dynal, Invitrogen) until the beads were drawn to the tube surface (~30 sec). The unbound fraction was collected, the

beads washed three times with each Buffer A, B and C, finally suspended in 200µl ChIP elution buffer and incubated at 65°C for 10 min to release bound fraction. This step was repeated once more with another 200µl of elution buffer and the two eluates were combined giving a 400µl final volume. In parallel 50µl of input sample was thawed and supplemented with 350µl of Elution buffer. To all samples, 21µl of 4M NaCl was added and the samples were left to incubate overnight at 65°C. Following incubation, 1µl of RNAase A (from 10 mg/ml, DNAaseI-free stock) was added and the samples were left to incubate at 37°C for 1 hour. This was followed by addition of 4 µl EDTA (from 0.5 M stock) and 2µl Proteinase K (from 10 mg/ml stock). The reaction was incubated for 2 hours at 42°C. Samples were cleaned up using Qiagen PCR clean up columns. IP and INPUT samples were eluted in 20µl 10 mM Tris (pH 7.5) before proceeding to PCR analysis.

2.18 Bicinchronic Acid Assay (BCA)

Protein concentration was measured using the BCA Protein Assay Kit (Pierce) according to the manufacturer instructions. The protein concentration was extrapolated for the Albumin (BSA) standard curve which was prepared in the range of 2mg /ml to 0.225mg/ml. The samples and standards (25µl) were incubated for 30 minutes at 37°C in BCA reagent, prepared according to manufacturer instructions. The protein concentration was determined by absorbance measurement on the Nanodrop 1000 (Thermo Scientific) BCA assay program settings.

2.19 CAT Assay

The Chloramphenicol acetyl transferase (CAT) assay is based on the imaging of ^{14}C acetylated forms of chloramphenicol. The CAT gene transfers the acetyl group from acetyl co-enzyme A to chloramphenicol. The relative differences in the resulting chloramphenicol products migrating down a thin layer chromatography plate are measured to determine the varying amount of CAT proteins present.

The cell samples prepared as described in Section 2.10.2. were lysed by addition of three cell volumes of 0.25mM Tris pH 8and three repeated freeze/thaw cycles (37°C for 5 minutes, vortexed, followed by snap freezing in dry ice). The cell suspension was centrifuged for 5 minutes at 13,000xg and the supernatant containing protein retained to a clean tube.

The protein concentration of each sample was determined using the BCA assay as described in previously. Typically, 5µg of the sample protein was prepared for the assay by addition of Tris pH 8.0 to 0.25M and then incubated at 60°C for 5 minutes. A mixture consisting of 0.5µl ^{14}C chloramphenicol, 2.5µl 40mm acetyl co-enzyme A and 17µl of 0.25M Tris pH 8 was added to the protein samples. Following 1 hour incubation at room temperature, the reaction was stopped by addition of 1ml ethyl acetate, vortexed and centrifuged for 5minutes at 13,000xg. Following centrifugation the upper solvent phase layer which contained chloramphenicol was removed to a fresh tube and left to evaporate O/N in a fume hood. 20µl of ethyl acetate was subsequently added to the dried samples and applied directly onto a silica gel TLC plate. The sample spots were placed 3 cm from the bottom of the plate with 1 to 1.5cm spacing between each sample. The sample spots were then air dried and the TLC sheet placed into a TLC containing 250ml of 95% chloroform/5% methanol. Migration was allowed to proceed until the solvent front was approximately 1 cm from the top of the plate. The plate was removed, dried in the fume hood and placed into a Phosphorimager cassette O/N. The Phosphorimager plate was scanned using a Fuji FLA 2000 Phosphorimager and the data evaluated by densitometry analysis using AIDA 2D software.

2.20 T cell Isolation and Stimulation

The anonymous single donor buffy coats obtained from the National Blood Service with ethical consent were used for primary quiescent T cell isolation. Histopaque (Invitrogen) was used to separate blood mononuclear cells (MNC) by density gradient, followed by a 560xg centrifugation for 30 minutes with no brake. The interphase containing the mononuclear cells was then washed twice with 30ml of PBS with 10% (v/v) FBS. A T cell negative isolation kit (Dynal,Invitrogen) was used to isolate quiescent T cells by negative selection using immunomagnetic beads,as specified in manufacturer instructions. The cells were counted using an improved Neubauer chamber and the 20µl of immunomagnetic beads conjugated to an antibody mix consisting of; anti CD-14 (depletes monocytes), HLA Class II DR/DP (to deplete B cells, activated T cells, neutrophils and monocytes) CD16 (a and b) (depletes NK cells), CD56 (activated T cells) and CD235 (erythroid cells) was added per $1x10^7$ cells. The cells were incubated with the mixture of antibodies for 20 minutes on a roller at 4°C, followed by centrifugation at 400xg for 8 minutes. The cells were then washed twice with 15ml of PBS supplemented with 2% (v/v) FBS. Cell pellets were resuspended in 50ml of PBS and incubated with anti-immunoglobulin conjugated beads for 15 minutes on a roller at room temperature. The cell bead complex was placed in a magnetic rack (Dynal, Invitrogen) and the supernatant containing the quiescent CD4+, CD8+ T cells transferred to a fresh tube. The collected cells were seeded in RPMI-1640 supplemented with 10% (v/v) FCS (Invitrogen) at 1-4x 10^6 cells/ml.

T cells were stimulated, when required by addition of CD3/CD28 magnetic beads (Dynal, Invitrogen) typically at 10µl of beads per $1x10^6$ cells). For the immunofluorescence experiments, the beads were dissociated from the cells with a magnet.

The T cell suspensions at $1x10^4$ were Cytospin were loaded into the cytospin cuvette containing the filter paper and lysine slides (as for IF procedure). The samples were centrifuged in a cytospin centrifuge at 200xg for 7 minutes, and processed as described in Sections 2.15.

2.21 Cell Cycle Analysis

Cell samples at $2x10^5$ cells were taken prior to and following stimulation and fixed in 400µl of 70% (v/v) EtOH at -20°C. Following centrifugation at 400xg for 8 minutes the pellet was resuspended in 400µl of FITC/PI cell cycle stain consisting of 40µg/ml PI (to stain for DNA content), 5µg/ml FITC (protein staining), and 1µg/ml RNAse1). The mixture was incubated at 37°C for 30 minutes and analysed on FACS Calibur flow cytometer. To analyse cell cycle data WinMDI 2.9 program was used to plot PI (FL2A) against FITC (FL-1H). Plotting area (FL2-A) against width (FL2-W) created a doublet discriminator gate to exclude doublets from the analysis. The cell cycle phase cell percentage was calculated by applying gates manually around cell populations in various cell cycle stages.

Chapter 3

Characterization of Zinc finger: methyltransferase fusion proteins

3.1 Introduction

3.1.1 Recombinant Zinc Finger Protein design, assembly, expression and purification

Four sites within the promoter region of *CDKN2B* were selected for methylation targeting studies. Prior to *in vivo* application I needed to determine that the ZF/Mtase components were functional in *in vitro* settings and could specifically bind to their intended target site. To do this, *in vitro* characterization of *CDKN2B* specific ZFs fused with prokaryotic methyltransferase M.HpaII was conducted.

What follows is the rationale for ZF selection and a description of their design, assembly, purification and an *in vitro* characterization of the affinity, binding specificity and enzymatic activity of the protein fusions used in this project.

To assess the *in vitro* binding specificity and targeted methylation activity of zinc finger methyltransferase fusions, the M.HpaII (F35H) prokaryotic methyltransferase was subcloned in frame 3' to various ZFs in a pGEX vector backbone, expressed and purified. The F35H mutation in M.HpaII has previously been shown to reduce the activity of the Mtase protein, allowing the zinc finger to dominate during targeted methyltransferase action [275].

The zinc finger and methyltransferase components of the construct were separated by a (Gly4Ser)3 linker, which allows a degree of independent function for each component of the targeted methyltransferase enzyme. The cloning strategy employed for the assembly of the GST constructs for purification is described in Figure 3-1 on the next page.

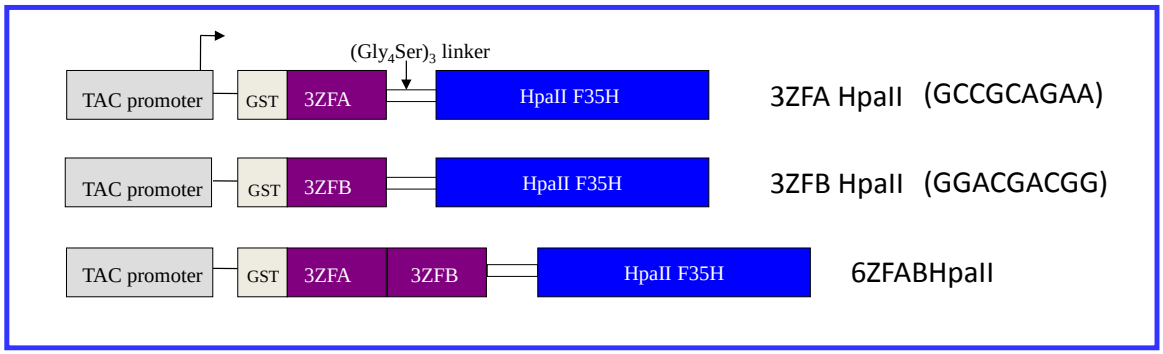


Figure 3-1 ZF/HpaII Cloning Strategy

The *HpaII* methyltransferase (F35H) gene was restricted with *BamHI/NdeI* and cloned into the multiple cloning site (MCS) of a pGEX plasmid vector already containing the three finger zinc finger proteins 3ZFA, 3ZFB, and six finger zinc finger protein 6ZFAB. Other commercially synthesized 6ZFs (PU1, MIZ, and MYC) were subcloned in the same manner. The expression of the cloned fragments under the control of the *TAC* promoter was induced by the lactose analogue isopropyl β-D thiogalactoside (IPTG). Additionally, pGEX vectors contain an internal *lacI*^{q} gene coding for a repressor protein involved in the regulation of expression through its ability to bind *TAC* promoter operator region, until induction by IPTG.

3.1.2 *CDKN2B* regulatory region

The rationale behind the selection of site biased ZFs to target *CDKN2B* promoter was based on extensive literature searches of what is known about this key cell cycle regulatory gene. Additionally, the Wisconscin GCG software package was used to search for TF binding sites in the *CDKN2B* regulatory region. The outcome of those searches is presented in Figure 3-2 which shows the key TFs with a role in *CDKN2B* regulation. *CDKN2B* (Genbank Accession Number ID AF513858) is a TATA- less promoter with an initiator sequence CCCCACTCT (Inr) [276]. The Regulatory Domain (RD) for the IN4/ARF locus lies in proximity to *CDKN2B*, and promotes locus repression via CDC6 [277]. This RD region was shown to be a region of Polycomb PRC2 (EZH2) and PRC1(BM1 and M33) localization and CDC6 interaction [278, 279]. The specific transcriptional regulators of *CDKN2B* includes signalling by TGF-β, which down regulates c-Myc *[280]*. Other factors that contribute to this TGF-β responsiveness are several Sp1 binding sites (consensus sequence GCCTCC, GGGGTGGGG and GGGGCGGAG) in proximity to the Inr element, and Miz-1 [281]. The Miz- 1 zinc finger protein is an essential interacting partner of the Myc: Max complex but can also bind to GFI-1 [282]. This interaction takes place in the region of the *CDKN2B* Inr and interferes with the recruitment of p300 to Miz-1, thereby preventing transcriptional activation of *CDKN2B*. The Myc: Max complex can also activate transcription via E – box elements. A potential E Box binding site within *CDKN2B* (TATGACAGGT) lays 420bp upstream from the Inr. Transforming growth factor-β (TGFβ) signalling is mediated through the SMAD family of signal transducer proteins. SMAD4 acts as a common mediator involved in the translocation of phosphorylated SMADs to the cell nucleus to effect cellular responses to TGF-β [283]. A potential SMAD3/SMAD4 site (consensus sequence CAGACAT) lays at around -400 from the initiator sequence. SNAIL1, SP1, and EGR1 were also demonstrated to be able to activate the *CDKN2B* promoter upon TPA stimulation [284]. Transcription factor PU.1 along with the interferon consensus sequence-binding protein/interferon regulatory factor 8 (ICSBP/IRF-8) was also shown to be able to bind to *CDKN2B* promoter upon IFNβ treatment [285]. A putative CCCTC binding site for a CTCF promoter insulator binding factor is located at the 5' end of the gene. Proposed functions for CTCF are in blocking enhancer activity and chromatin insulation.

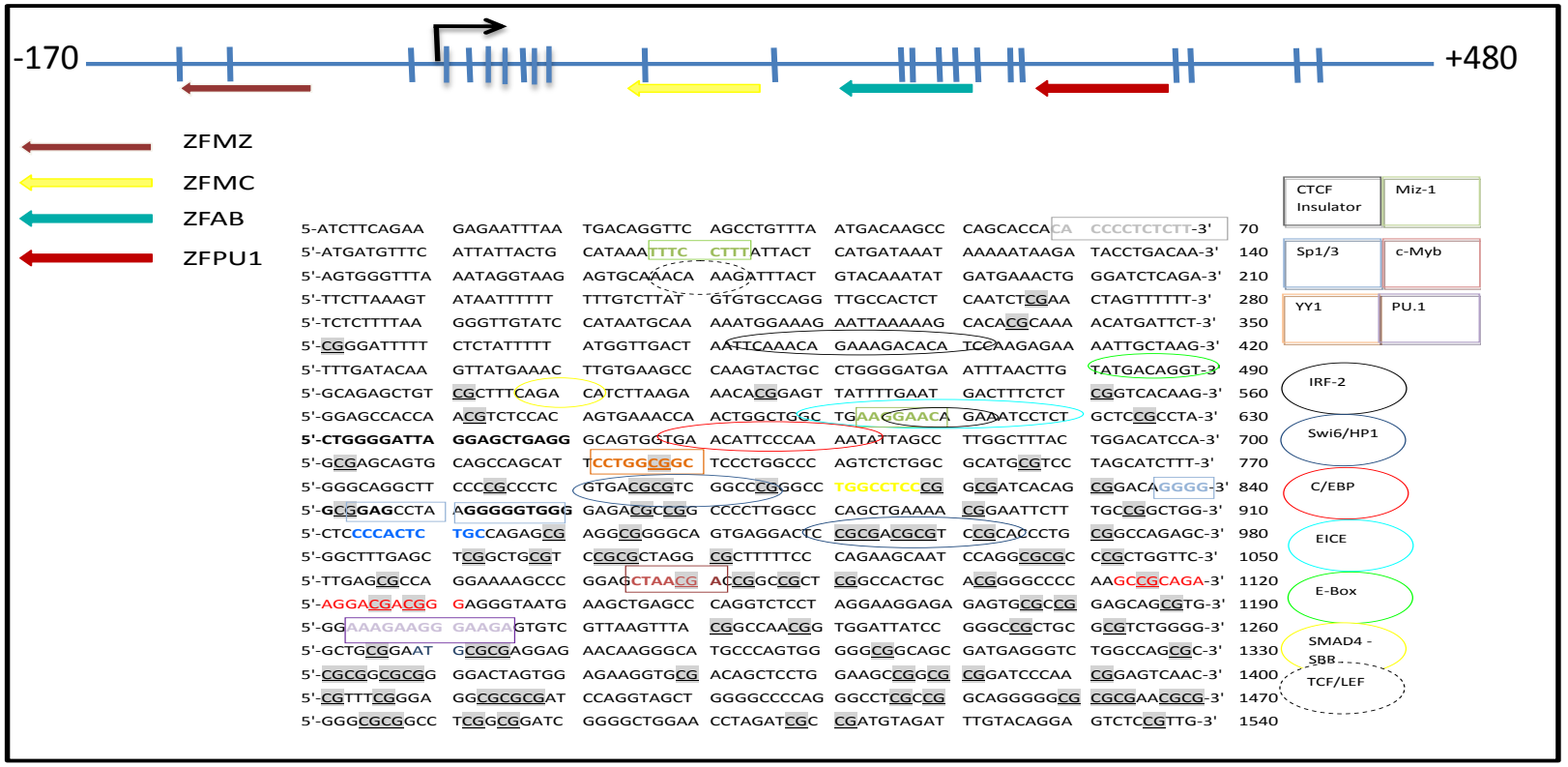


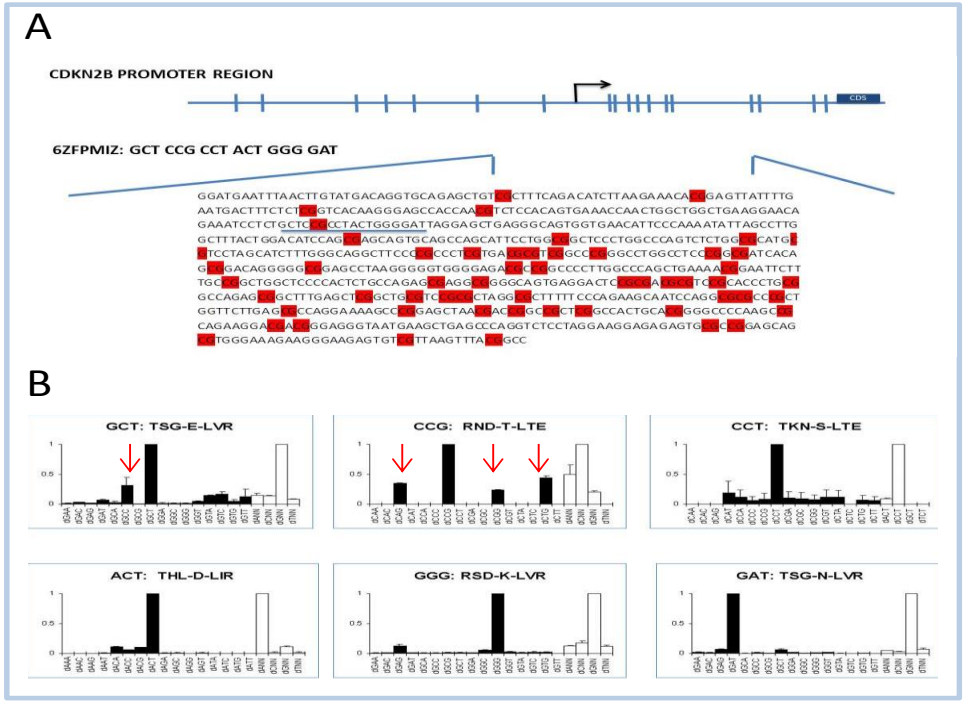
Figure 3-2 *CDKN2B* promoter regulatory region.

Key TF sites, Transcription Start Site, zinc finger binding and methylation sensitive sites highlighted. YY1 consensus sequence CCTGGCGGC -180 from Inr and c-Myb CTAACGA at + 145 for Inr, are known to be sensitive to methylation of their binding sites, [286, 287], while TFs such as Sp1 are thought to be sterically blocked by binding of methylcytosine binding proteins to its methylated binding site. Inr sequence is highlighted in blue text which will be discussed later on in the Chapter. The scale of the blue line at the top of the diagram representing the *CDKN2B* CpG and the target sites of the ZFPs is 650bp, starting at -170bp to +480bp from TSS.

The zinc finger proteins were designed using the software Zinc Finger Tools, which is available freely through an organization established to promote zinc finger engineering called the Zinc Finger Consortium, designed by Barbas' laboratory [260] (www.zincfinger.org). The modular approach and Base function of the Zinc Finger Tools software ensures that each zinc finger targeting a triplet code is assigned a score [288]. This score predicts a number of 'off- target' sites that a zinc finger might recognise and is calculated from ELISA specificity graphs. As a general rule of thumb, the higher scores generated by Base function correlate with the lower number of 'off- target' sites that a zinc finger recognises. ZFs were commercially synthesized with the exception of one which was assembled in the lab using the SOE-PCR approach. Figure 3.3A-C shows the process of construction of all the *CDKN2B* specific six finger ZFs described in this thesis. Figure 3-3D describes the process of SOE-PCR 6ZFAB assembly process.

Designed zinc finger proteins contain two N terminal beta strands (YKCPECGKSFS) and an alpha helix C terminal backbone (HQRTH) which is a region not involved in DNA recognition and contact. Zinc finger modules are joined via fixed sequences at the N (LEPGEKP) and C termini (TGKKTS) to form a TGEKP linker.

The 6ZFAB was generated by fusion of two three zinc finger modules to target consecutive *CDKN2B* promoter regions. Each single zinc finger module recognizes three base pairs of DNA, which means that following the assembly, zinc finger proteins can target an 18 (six zinc finger) base pair contiguous DNA sequence. The oligonucleotides coding for the beta sheet and alpha helix modules were synthesized, assembled using the Splicing by Overlapping Extension (SOE-PCR) approach, and cloned into pGEX vectors for expression in *E.Coli* and GST purification. The Zinc finger array was selected to bind to an 18bp sequence 5'-GCCGCAGAAGGACGACGG-3'. The site is located +180 from the *CDKN2B* Transcription Initiator element Inr [289].



C

6ZFPMIZ Target Site	GCT CCG CCT ACT GGG GAT
N-terminal backbone	YKCPECGKSFS
C-terminal backbone	HQRTH
ZF linker	TGEKP
C-terminal fixed	TGKKTS
N-terminal fixed	LEPGEKP
$(Gly_4Ser)_3$ linker	GGGGSGGGGSGGGGS

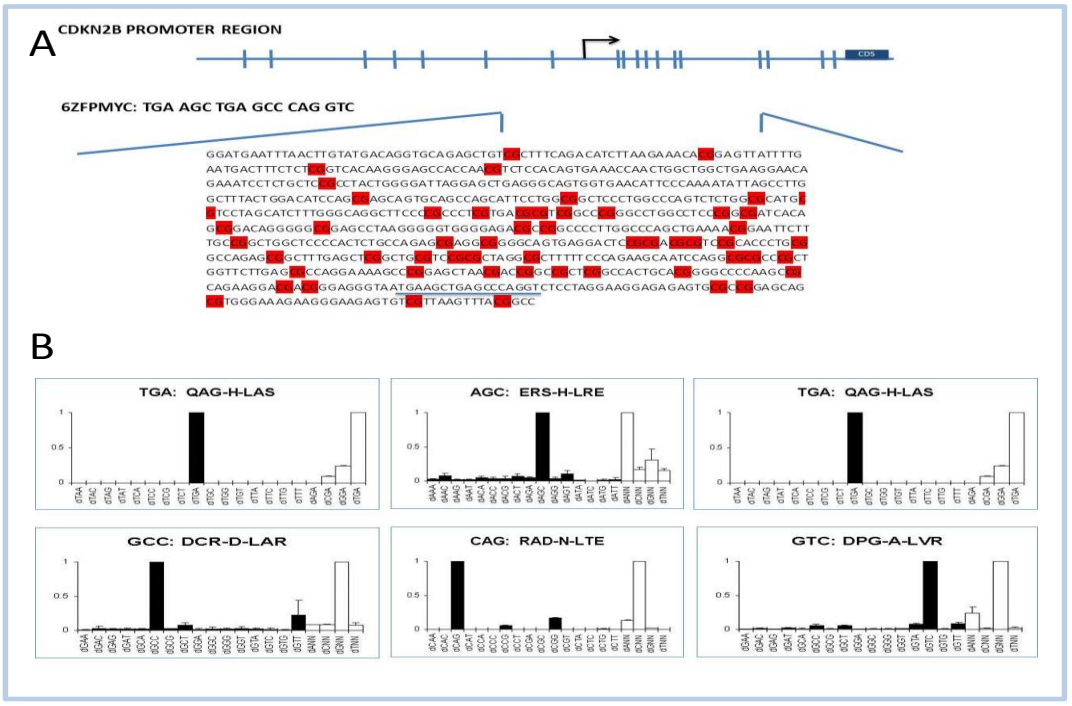
Finger Module	Triplet	Helix
1	GAT	TSGELVR
2	GGG	RNDTLTE
3	ACT	TKNSLTE
4	CCT	THLDLIR
5	CCG	RSDKLVR
6	GCT	TSGNLVR

The amino acid sequence of the ZFPMIZ predicted to bind to the target sequence

LEPGEKPYKCPECGKSFSTSGNLVRHQRTHTGEKPYKCPECGKSFSRSDKLVRHQRTHTGEKP YKCPECGKSFSTHLDLIRHQRTHTGEKPYKCPECGKSFSTKNSLTEHQRTHTGEKPYKCPECGK SFSRNDTLTEHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGKKTSGGGGSGGGGSGG GGS

Figure 3-3A. Construction of *CDKN2B* specific zinc finger arrays using Zinc Finger Tools.

A *CDKN2B* promoter region with the location of 6ZFMIZ recognition region (underlined blue). The amino acid sequence of the zinc finger array predicted to bind to this *CDKN2B* promoter target sequence was generated by entering the target site of interest into the Zinc Finger Tools software. For 6ZFMIZ the target site was 5' GCT CCG CCT ACT GGG GAT 3'. **B** The modular zinc finger best able to bind each target triplet code was generated by the software using a scoring system based on base function. This scoring system uses ELISA specificity graphs to predict a number of off -target sites that a zinc finger may recognise, with the highest score assigned to the fewest number of the off- target sites. 6ZFMIZ has several off-target sites (arrowed). **C** All of the zinc finger arrays designed and constructed for this project used an identical N-terminal and C-terminal backbone, both based on the human transcription factor Sp1, TGEKP linker, N-terminal fixed, and C-terminal fixed frame.



C

6ZFPMYC Target Site	TGA AGC TGA GCC CAG GTC
N-terminal backbone	YKCPECGKSFS
C-terminal backbone	HQRTH
ZF linker	TGEKP
C-terminal fixed	TGKKTS
N-terminal fixed	LEPGEKP
$(Gly_4Ser)_3$ linker	GGGGSGGGGSGGGGS

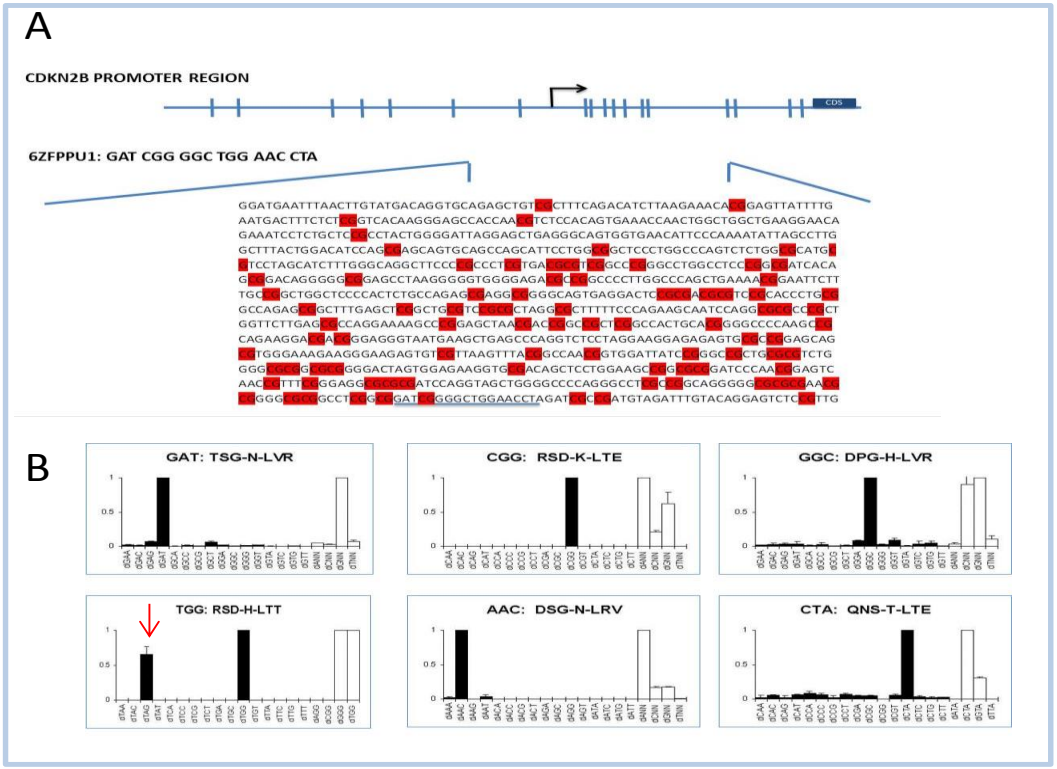
Finger Module	Triplet	Helix
1	GTC	QAGHLAS
2	CAG	ERSHLRE
3	GCC	QAGHLAS
4	TGA	DCRDLAR
5	AGC	RADNLTE
6	TGA	DPGALVR

The amino acid sequence of the ZFPMYC predicted to bind to the target sequence

LEPGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCPECGKSFSRADNLTEHQRTHTGEK
PYKCPECGKSFSDCRDLARHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGEKPYKCPE
CGKSFSERSHLREHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGKKTS
GGGGSGGGGSGGGGS

Figure 3-3B. Construction of *CDKN2B* specific zinc finger arrays using Zinc Finger Tools.

A The location of the 6ZFMYC binding site in relation to the *CDKN2B* gene is shown highlighted in blue and the location of each CpG site is shown in red. The target site 5' TGAAGCTGAGCCCAGGTC 3' for 6ZFMYC was entered into the Zinc Finger Tools software to generate the amino acid sequence of the zinc finger array predicted to bind to the *CDKN2B.* **B** The ELISA specificity graphs were generated using a scoring system utilizing base functions as already explained in Figure 3-3A. **C** The general modular composition of the 6ZFMYC is shown in the table.



C

6ZFPPU1Target Site	GAT CGG GGC TGG AAC CTA
N-terminal backbone	YKCPECGKSFS
C-terminal backbone	HQRTH
ZF linker	TGEKP
C-terminal fixed	TGKKTS
N-terminal fixed	LEPGEKP
$(Gly_4Ser)_3$ linker	GGGGSGGGGSGGGGS

Finger Module	Triplet	Helix
1	CTA	TSGNLVR
2	AAC	RSDKLTE
3	TGG	DPGHLVR
4	GGC	RSDHLTT
5	CGG	DSGNLRV
6	GAT	QNSTLTE

The amino acid sequence of the ZFPPU1 predicted to bind to the target sequence

LEPGEKPYKCPECGKSFSQNSTLTEHQRTHTGEKPYKCPECGKSFSDSGNLRVHQRTHTGEKP
YKCPECGKSFSRSDHLTTHQRTHTGEKPYKCPECGKSFSDPGHLVRHQRTHTGEKPYKCPEC
GKSFSRSDKLTEHQRTHTGEKPYKCPECGKSFSTSGNLVRHQRTHTGKKTS
GGGGSGGGGSGGGGS

Figure 3-3C. Construction of *CDKN2B* specific zinc finger arrays using Zinc Finger Tools.

A The location of the 6ZFPU1 binding site in relation to the *CDKN2B* gene is shown highlighted in blue and the location of CpG sites is shown in red. The amino acid sequence of the zinc finger array corresponding to the *CDKN2B* target sequence was generated by entering the target site of interest into the Zinc Finger Tools software. For 6ZFPU1 the target site was 5' GAT CGG GGC TGG AAC CTA 3'. **B** The ELISA specificity graphs were generated using a scoring system based on a base function and the off-target site is arrowed as already explained in Figure 3-3A. **C** The general modular composition of the 6ZFPU1 is shown in the table.

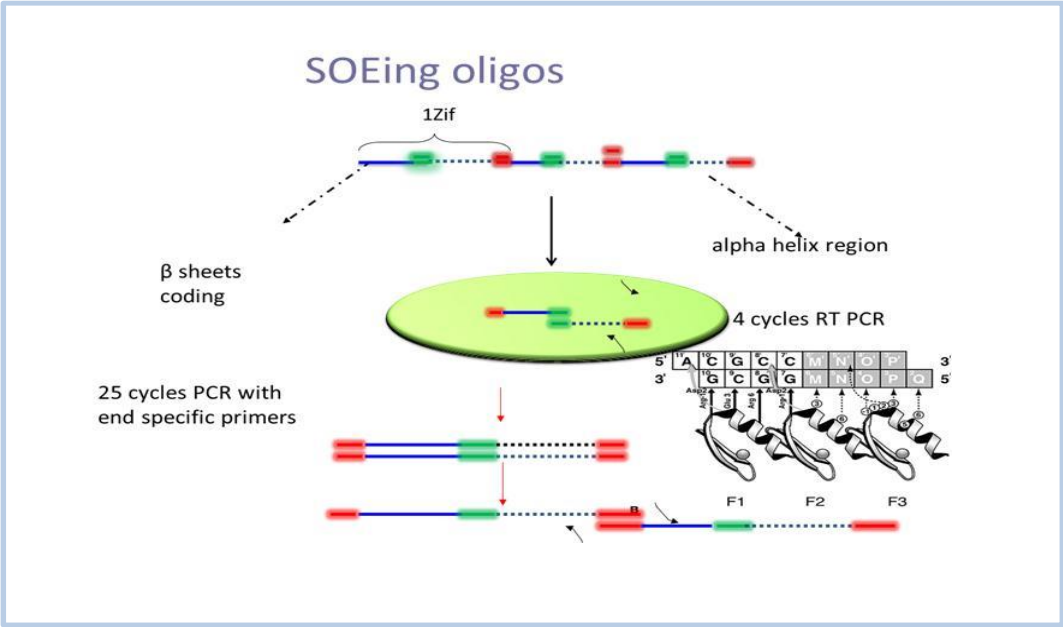


Figure 3-3D. Construction of *CDKN2B* specific 6ZFAB using SOE-PCR approach.

Zinc finger proteins were sequentially assembled by PCR overlap to generate two 3ZFmodules 3ZFA and 3ZFB. The oligonucleotides were designed to contain 15-20 bp overlaps to allow the complementary fragments to anneal. The template sequence was then filled in using DNA polymerase to create longer DNA fragments [290]. This was followed by a standard PCR reaction using primers specific for sequences in the extreme 5' and 3' ends of the assembled gene, leading to PCR product amplification. The primers had 5' NdeI and 3' *HindIII* overhangs for subsequent cloning into pGEX vector.

Following SOE-PCR assembly the gene products for the two different zinc finger arrays (designated ZFA and ZFB) were gel purified (Figure 3-4A) , TOPO cloned and sequenced.

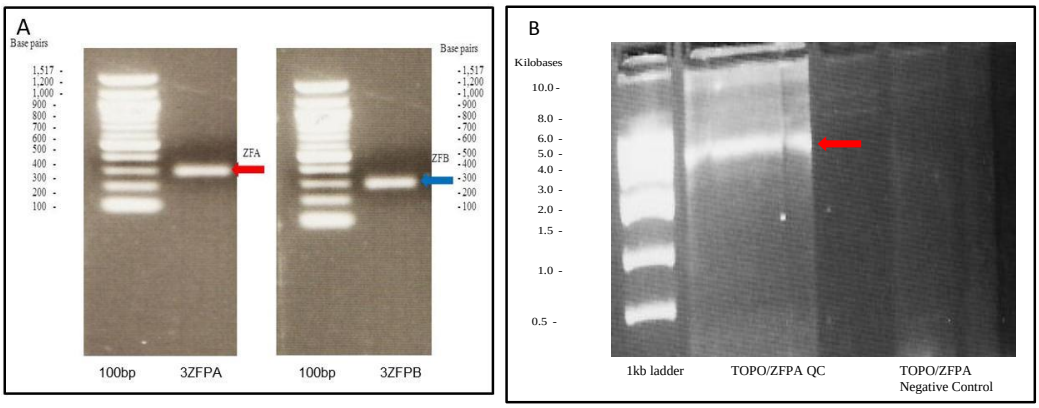


Figure 3-4. Gel purification and QC mutagenesis of 3ZF Arrays.

Panel **A** shows gel resolution of two three ZF arrays ZFA and ZFB which are 288bp PCR products. Panel **B** Quick-change mutagenesis products were generated for both TOPO/ZFA shown here and TOPO/ZF not shown. Negative control was TOPO/ZFA subjected to the quick-change reaction in the absence of Phusion enzyme.

Sequencing data revealed a consistent single base insertion in the β sheet region CATATCCG**C**CATCC (ABI File sequencing data in Appendix Section 1 CD-ROM) for ZFA and a single base deletion in the ZFB α helical region ATTTAAC**A**AGG. Quick-change mutagenesis was used to correct for these errors. Self- complementary oligonucleotide primers were synthesized using the Quick-Change protocol described in the Methods Section 2.9.1. The primers containing the desired mutation were designed to anneal to both strands of the TOPO/ZFA or ZFB plasmid at the same region. Primers were extended during temperature cycling with Phusion DNA polymerase PCR and the reaction treated with *R.DpnI* endonuclease. This enzyme only cuts *Dam+* methylated DNA (5´-Gm6ATC-3´) and therefore digests only parental DNA template, which was isolated from (Top10) *Dam+ E.Coli* cells (Figure 3-4B).

3.1.3 Recombinant protein purification

NEB-Express competent cells (genotype: *fhuA2 [lon] ompT gal sulA11 R (mcr-73 mini Tn10- - Tet$^{s)\ 2}$ [dcm] R(zgb-210..Tn10- - Tet^{s}) endA Δ(mcrC-mrr)114.IS10)* were transformed with the pGEX AB or AB plasmid. A single colony was used to inoculate 25ml of LB medium containing ampicillin. The culture was grown overnight at 37°C, and the cells were collected by centrifugation at 2000xg for 10 minutes at 4°C. The cell pellet was resuspended in 1ml of LB media, and used to inoculate 250ml LB media supplemented with 100µl 1M $ZnCl_2$, glucose 0.1% (w/v) and 250µl ampicillin and incubated at 37°C until O.D.600 nm reached 0.8 (around 2 hours). 5ml of culture was retained for use as an un-induced control. IPTG (0.5mM) was then added to the remainder of the cell solution, for a further 2 hour incubation at 30°C. Following incubation, the cells were pelleted by centrifugation at 5000xg for 15 minutes at 4°C, and the pellet was washed in 25ml PBS on ice by pipetting up and down gently to disperse clumps. The cell suspension was then centrifuged at 5000xg for 10 minutes at 4°C and the pellet was either frozen at -70°C for subsequent purification or processed immediately. The cell pellet was resuspended in 10ml GST resuspension RP1 buffer. The protein suspension was then sonicated for 30 seconds twice at 60Hz, followed by the addition of 1ml of Triton-x-100 10% (v/v) to the suspension and left for 30 minutes on a roller at 4°C. During this incubation, GST-sepharose beads were pre-washed with 10ml RP1 buffer. Following 30 minutes incubation, the protein suspension was centrifuged at 3500xg for 10 minutes at 4°C, and the supernatant was transferred to the GST-sepharose beads for a further 30 minutes roller incubation at 4°C. The GST beads: protein complex suspension was centrifuged at 900xg for 5 minutes and the supernatant decanted. The beads/protein complex was then washed four times with 50mM Tris-HCl. The GST beads were then resuspended in 1ml of 50mM Tris-HCl, supplemented with 20% glycerol (v/v) and transferred onto a 5ml disposable column (Pierce). The column was washed with 2ml of 50mM Tris-HCl 20% glycerol (v/v), and the protein was then eluted with 5ml of reduced glutathione RP2 buffer. Five 1ml fractions were collected and frozen at -80°C. Purified GST tagged protein was run on 10% SDS-PAG to detect GST tagged eluted zinc finger protein using Coomassie blue staining shown in Figure 3-5.

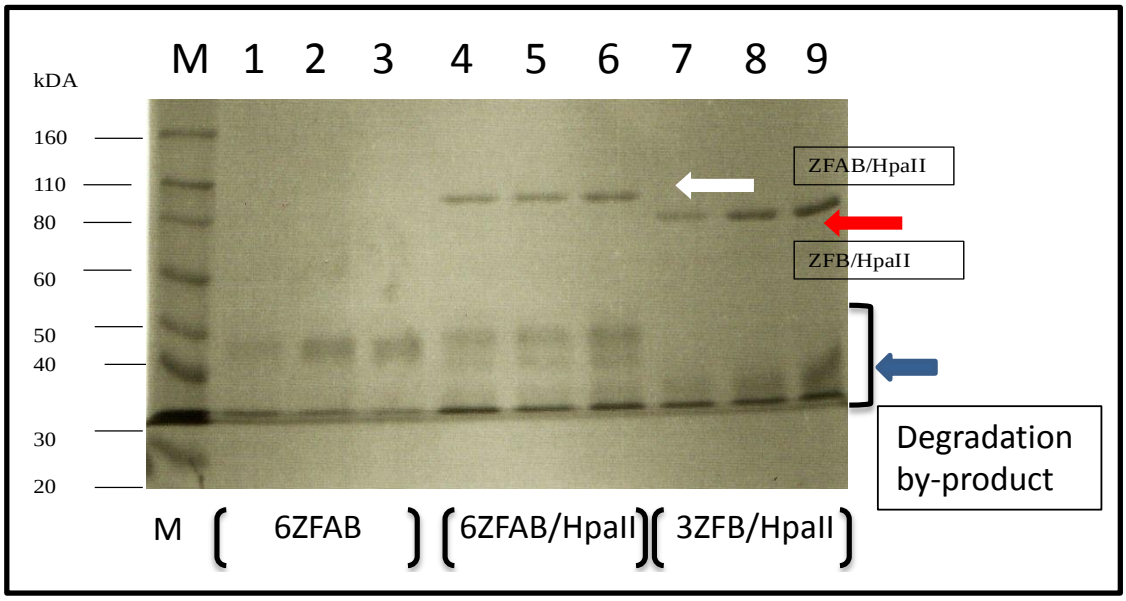


Figure 3-5 SDS PAGE Analysis of GST purified ZF-Mtase

To determine the quality of the expressed and GST purified protein, the proteins were analysed by SDS PAGE. Figure shows GST 6AB Fractions 1-3 (lanes1-3), 6ZFAB/HpaII Fractions 1- 3 (lanes 4-6) and 3ZFB/HpaII Fractions 1-3 (lanes 7-9) eluates. Empty pGEX- 5MCS cloning vector produces a 24kDa GST protein, while the six zinc finger protein ZFAB is 22kDa and HpaII is ~ 40kDa, so the expected size for 6ZFAB/HpaII fusion protein is ~ 86kDa. The expected size for 3ZFB/HpaII is ~ 75 kDa. Novex ® Sharp Protein Standard was used for molecular weight estimation.

3.2 Measurement of ZF/Methyltransferase binding affinity using EMSA

Six finger proteins were chosen to target a single unique site within the mammalian genome and initially their binding affinity and specificity was evaluated using Electrophoretic Mobility Shift Assay (EMSA). Using EMSA it is possible to determine the binding affinity and specificity of a protein for its target DNA sequence, as well as the binding kinetics of those interactions [291, 292]. Binding kinetics analysis depends upon the law of mass action which is based on several assumptions. Firstly, that all of the binding sites on a DNA sequence are equally accessible to the binding protein. Secondly, only free or bound interactions are accounted for whilst partial binding interactions are not allowed by this model. Binding is a result of a collision event occurring by diffusion between the zinc finger/methyltransferase and the target DNA which once bound remain so for a random amount of time. When the rate at which new ZF/Methyltransferase : DNA complexes are formed (k_{on} rate) equals the rate at which the ZF/Methyltransferase : DNA complexes dissociate (k_{off} rate), equilibrium is reached. The equilibrium dissociation constant K_d can then be defined as follows:

$$\frac{\text{Reactants}}{\text{Products}} = \frac{[\text{DNA}][\text{ZFMt}]}{[\text{DNA:ZFMt}]}$$

Receptors with a high affinity for a particular ligand will have low K_d values, because a low concentration of ligand will be required to bind half the receptors. The assay is conducted routinely in our lab and the previous work using a time course for six finger proteins has shown that the equilibrium for the binding reaction to occur is reached following 10 minutes incubation period with the target DNA duplex oligonucleotide radiolabelled with ^{32}P.

Four six zinc finger protein (PU1, MYC, MIZ and JMJ) HpaII methyltransferase fusions were subcloned into pGEX vectors, expressed in *E.Coli* and purified as described in 3.1.1 and 3.1.2. Six finger proteins are capable of recognizing a contiguous 18bp sequence, considered statistically unique and highly unlikely to occur at random when looking at the context of the human genome of around $4x10^{18}$ bases in size [293]. The expression profile of purified GST:ZF/HpaII protein fractions is shown in Figure 3-6 A and B.

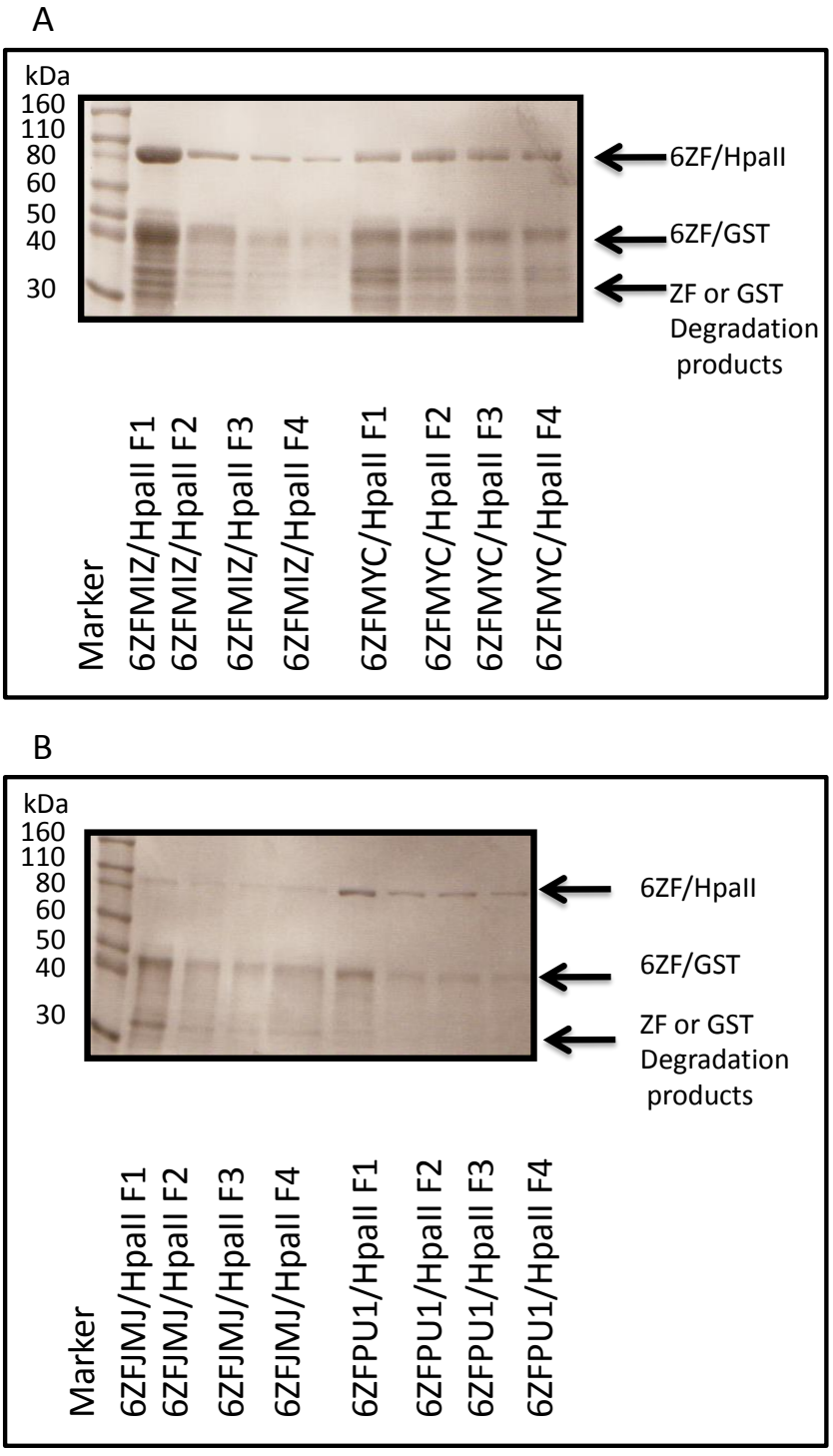


Figure 3-6 SDS PAGE Analysis of GST purified 6ZF/HpaII fractions

Four different GST purified 6ZF/Mtases Fractions 1-4; 6ZFMIZ/HpaII, 6ZFMYC/HpaII, 6ZFJMJ/HpaII and 6ZFPU1/HpaII were resolved on a 10% SDS PAG to determine the quality of the purified protein fractions. Novex ® Sharp Standard Protein Marker was used as a reference point. Six zinc finger protein 6ZF/HpaII fusions expected size was 86kDa.

The quality and purity of the purified protein fractions were evaluated using SDS gel electrophoresis. In Figure 3.5. 6ZFAB/HpaII purified protein is seen as a dominant band corresponding to the predicted molecular weight of 84kDa. Faint, lower molecular weight bands, likely corresponding to products of degradation are also seen. In Figure 3.6. all six finger proteins are confirmed to be of the identical size and approximately equivalent purity. Lower molecular weight bands that can be observed for all those proteins could be attributed to either degradation or protease cleavage of the purified protein resulting from purification process. However, addition of protease inhibitors was standard practice during protein isolation and purification. These low molecular weight bands are approximately 26-40kDa in size and correlate with; GST (26kDa) or GST tagged six finger ZFs (48kDa) degradation by-products. It is however, unlikely that any of the observed by- products contain methyltransferase component. This is because the 6ZF/Methyltransferase proteins are expressed and purified as N terminal – GST – ZF – fusions. Purification of the protein using GST beads is unlikely to capture any of the 6 potentially degraded methyltransferase which lack the N- terminal GST – ZF component. Protein dimerization during purification is however a possible explanation for the observed additional bands. Similar, observations in the lab were made previously during the purification process of ZF-Mtase-GST fusions, and the resulting by-products were confirmed by western blot with anti GST antibody to be zinc finger/GST fusions resulting from proteolytic cleavage.

Purified protein fractions visualized using denaturing SDS electrophoresis with the least degradation, or cleavage by- products were used for subsequent EMSAs (Figure 3.7). Purified protein was more accurately quantitated using the Bicinchronic Acid Assay (BCA) (for assay details see Methods Chapter 2 Section 2.18). Briefly, the BCA was used to determine quantitatively the concentration of a protein calibrated against the Albumin Standard at 2mg/ml. The assay is based on a colorimetric measurement of the intensity of purple product formed as a result of BCA and Cu^{+} reaction under alkaline conditions. Protein present in the solution is involved in a reduction of Cu^{2+} to Cu^{+}, so there is a direct correlation between the concentration of a protein, presence of Cu^{+} and the intensity of the purple coloured reaction product as measured by a spectrophotometer. For ZFAB/HpaII protein fusion, fraction 1 and for all the other HpaII fusion proteins, fractions 4 were used for all of the subsequent analyses. Protein fractions were incubated with DNA duplex oligonucleotides containing ZF target sequences

labelled with ^{32}P as described in Methods Chapter 2.11. This section also describes EMSA assay set up and visualization of protein-DNA bound complexes.

Previous work in our lab has shown that the binding affinity of the ZFs increases with the number of zinc fingers, so that by increasing the number of modules from three to four for example leads to an approximate 2-3 fold increase in binding affinity of that ZF for its target site. Using comparative protein levels, the upper limit binding affinity for each of the five six finger ZFs was evaluated for target and non-specific DNA (Figure 3.7). Although SDS PAGE analysis showed intact proteins of the correct size and of comparable quality to all other ZFs analysed, 6ZFMYC/HpaII did not bind at all to its target sequence. A possible explanation is that this protein has not folded properly in bacterial cells. All the other ZFs bind to their respective target sequences, but with different binding affinities. 6ZFMIZ/HpaII binds to its recognition site with the lowest apparent affinity and 6ZFPU1/HpaII with the highest apparent affinity. For 6ZFAB/HpaII a clear band signifying protein binding appears at lower protein concentration compared to 6ZFMIZ/HpaII. The 6ZFAB/HpaII contains the only zinc finger component not synthesized commercially. It was assembled from designed oligonucleotides in the lab by the SOE-PCR method and subsequent fusion of the two individual 3 finger ZFs. Following its assembly, sequencing and purification it became clear that during the oligonucleotide synthesis one amino acid was omitted which resulted in a compromised, though functional final module for each of the three finger proteins.

For each of the ZFs used in EMSAs, controls using DNA with HpaII sites but lacking the ZF target sequence show the appearance of DNA/protein complex bands only at the highest protein concentrations. This suggests that there exists a low level of non-targeted binding by the methyltransferase component of the ZF/HpaII fusion protein binding to its intrinsic HpaII binding site which is present of the non-specific probe too. This is in line with the data from previous work in our laboratory on four finger ZFs confirming that the HpaII methyltransferase used in ZF fusions has an intrinsic binding affinity for its binding site [294]. The overall binding affinity of the ZF/HpaII fusion is significantly higher than the binding affinity of ZF alone.

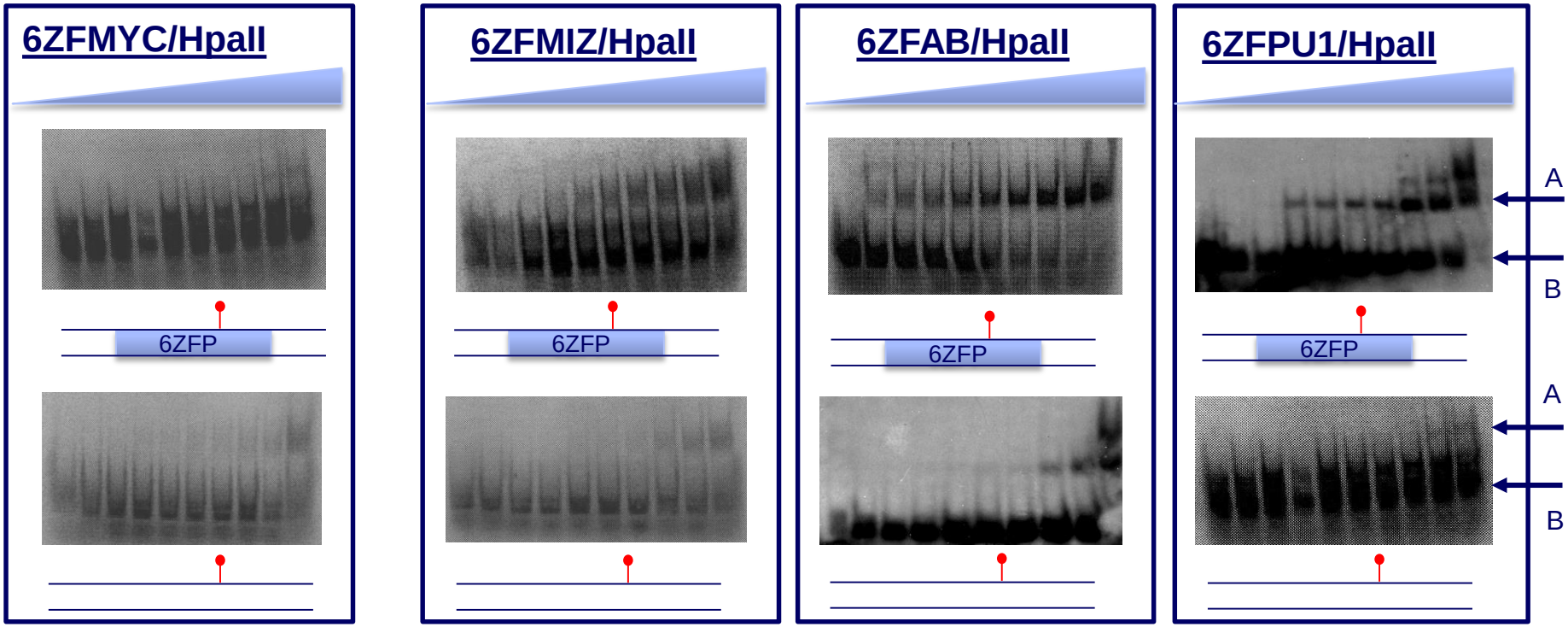


Figure 3-7 EMSA assay of 6ZF/HpaII proteins.

EMSA analysis of four different 6ZFs; 6ZFMYC/HpaII, 6ZFMIZ/HpaII, 6ZFAB/HpaII and 6ZFPU1/HpaII binding to ZF/HpaII binding site (top panels) and ZF negative site (bottom). The first lane on each slide contains no protein and the range for all of the 6ZFs used in these EMSAs is 1pmol to 0.6nmol. ^{32}P labelled oligonucleotide complexes with or without ZF target site (blue box representing 18bp target sequence for a given ZF) and/or containing a HpaII site (red lollipop) were incubated with their respective ZF. When the ZFs (with the exception of 6ZFMYC/HpaII which did not bind at all) were incubated with the oligonucleotide containing its target sequence and HpaII site, bound protein/DNA complexes were observed arrowed A and their intensity increases with the amount of ZF-Mtase protein. ZFs incubated with the oligonucleotides containing just the HpaII site reveals protein/DNA bound complexes only at the very highest protein concentrations. This indicates high specificity of the 6ZF to its target sequence.

The K_d values were calculated from two sets of independent gel shifts. The results in Figure 3.8 show the difference in K_d values between different ZFs. It was important to gain insight into relative binding affinities of these 6ZFS as a prelude for studies in the endogenous setting when targeted to the *CDKN2B* gene.

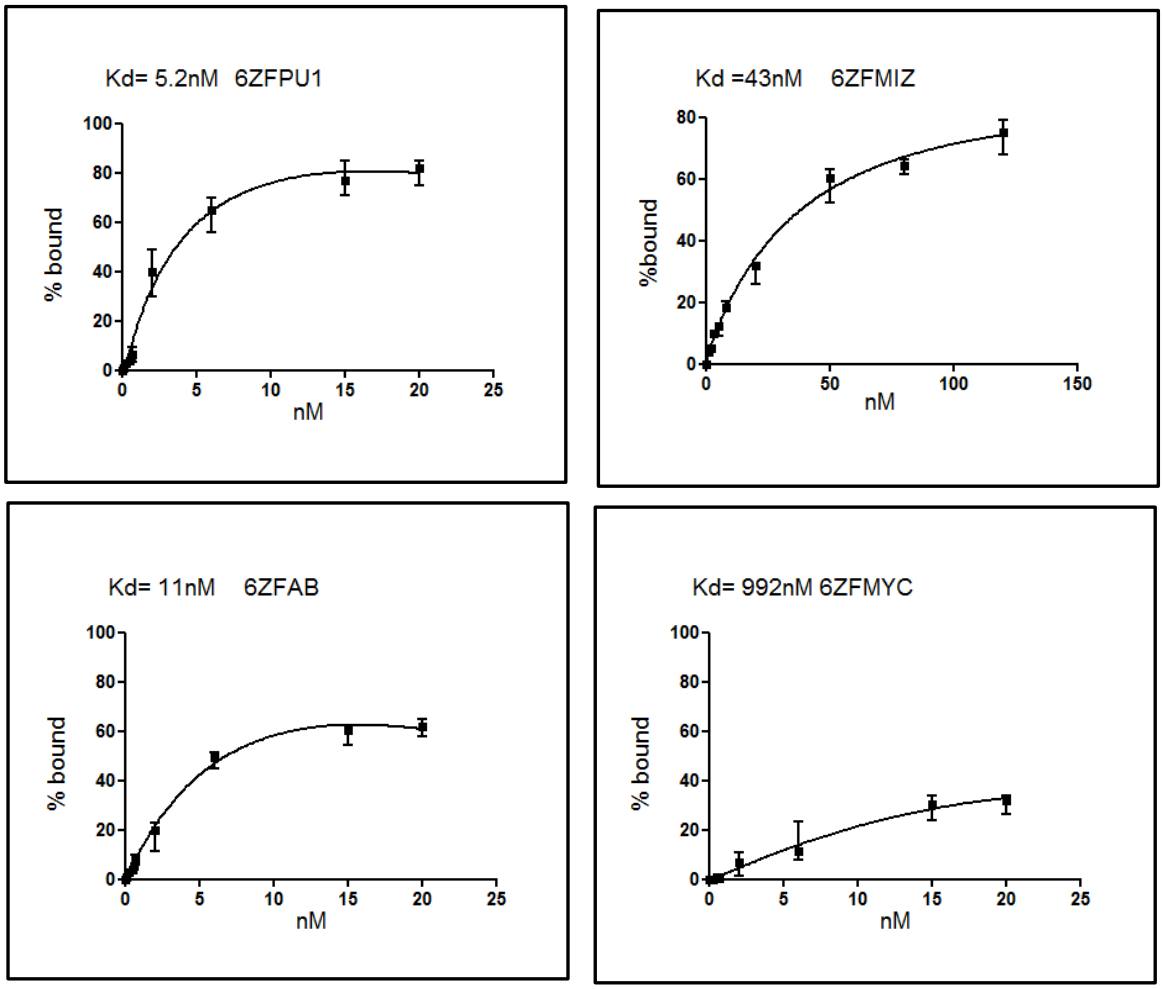


Figure 3-8 Binding curves and K_d values generated for each of the 6ZFs.

EMSA data was analysed using Graph Pad prism software. Four different 6ZFs; 6ZFMYC/HpaII, 6ZFMIZ/HpaII, 6ZFAB/HpaII and 6ZFPU1/HpaII targeted to the *CDKN2B* promoter were examined. The concentration range for the 6ZFs used was from 1pmol to 0.6nmol. Binding of 6ZF to the radioactively labelled oligonucleotide containing the target sequence is represented by the graph. Prism – GraphPad Software was used to quantify the binding affinity for each of the ZFs. The K_d values, and binding curves were generated, using a non- linear regression (curve fit), based on one-site binding (hyperbola) to generate a best fit curve (y axis = % protein bound). The K_d (nM) is calculated as the protein concentration at which 50% of the protein is bound and varied significantly between the four 6ZFs used from Kd 5.2nM for 6ZFPU1, 43nM for 6ZFMIZ, 11nM for 6ZFAB, and 992nM for 6ZFMYC.

3.3 *In vitro* methylation activity assay

In order to evaluate the catalytic activity of methyltransferase fusion proteins, *ex vivo* assays were employed to look at *in vitro* targeted methylation using Litmus™ 28i vectors. Having confirmed the binding affinity of the 6ZF/HpaII fusions, the same protein fractions used for EMSAs were employed to determine the intrinsic levels of DNA methyltransferase activity.

Initially, a crude estimate of methyltransferase activity was obtained from a restriction analysis of the coding plasmids for one of the ZFs, 6ZFAB fused with the CpG methyltransferase SssI. The expression plasmids coding for 6ZFAB/SssI were expressed and purified from *E.Coli* (as described previously). The plasmid was then restricted for 4 hours using enzymes *R. HpaII* or *R.MspI* and electrophoresed on an agarose gel. Restriction enzyme *R.MspI* is only sensitive to methylation at the first cytosine of its recognition sequence (5'CCGG3') which it cuts at CG. On the other hand, *R. HpaII* is sensitive to methylation at the second cytosine (5'CCGG3'), and although it restricts an identical sequence to *R.MspI*, it does so in a methylation dependant manner. So if the CG within the 5'CCGG3' recognition pattern is methylated then cleavage by the *R.HpaII* enzyme will be blocked. This principle therefore, allowed for expression vectors to be scrutinized for the level of 'self'- methylation. Products of higher molecular weight were observed when *R.HpaII* restriction was blocked as a consequence of methylation as shown in Figure 3-9 A for 6ZFAB/SssI. This allows for a crude assessment of the methyltransferase activity of the fusion proteins coded by the plasmid. The results show that the prokaryotic methyltransferase SssI is enzymatically active as evident by the presence of high molecular weight bands resulting from a high level of cytosine methylation.

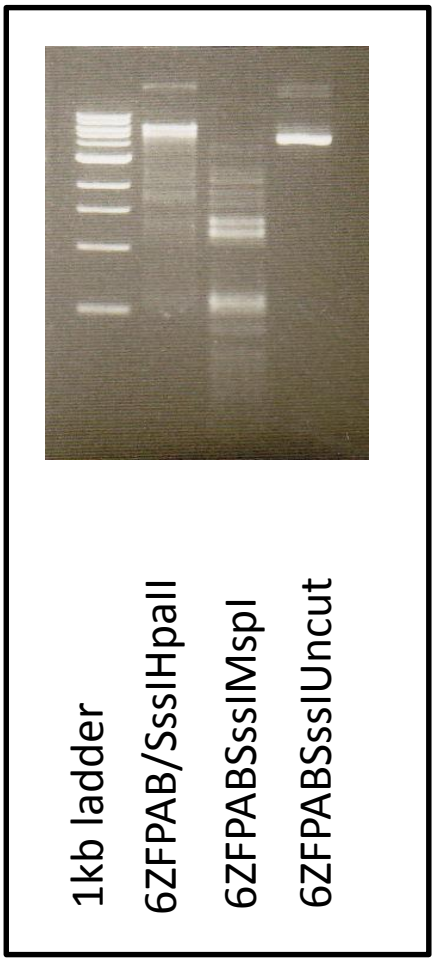


Figure 3-9. SssI methyltransferase activity assessment using methylation sensitive restriction analysis.

6ZFPU1/HpaII and 6ZFPU1/SssI plasmids were restricted for 4 hours with either *R.MspI*(M) or *R.HpaII*(H) restriction enzymes. Different restriction pattern profiles can be seen when the plasmids are restricted with *R.HpaII* enzyme which correspond to high molecular weight products arrowed.

Plasmid restriction demonstrates variable methyltransferase activity of all ZFAB/SssI constructs. In order to investigate in more detail specific targeted methylation by these constructs the *ex vivo* binding assay previously developed by our lab was employed. Oligonucleotides with (LZF) or without (LNS) zinc finger specific sites were synthesised and cloned into the Litmus™ 28i target vector as *XbaI/BglII* fragments. The vectors contain 22 HpaII sites uniformly distributed through the 3kb sequence. All of the 6ZF/HpaII fusion proteins (same fractions as used in EMSAs) as well as the 6ZFAB/SssI fusions were used to target HpaII/CpG sites within the target Litmus™ 28i vector. The ability of zinc finger protein methyltransferase fusions to selectively distinguish between the LZF and LNS target vectors and the resulting methylation patterns were assessed via subsequent restriction of plasmids with R. HpaII restriction enzyme. Detailed assay set up is described in Methods Section 2.12. Targeting of ZF/methyltransferases involved incubation of purified protein with target plasmids, followed by a restriction digest for analysis of targeted methylation. The restriction analysis is based on the methylation sensitivity pattern of *R.HpaII* restriction enzyme between target and non-specific plasmid substrates. A unique restriction pattern should be generated following successful targeted methylation shown in Figure 3.-10A. Titration of protein with the target plasmid should lead to a comparable increase in the appearance of high molecular weight restriction products. Figure 3-10C shows a selection of results obtained for the fusion proteins. As the results show it was not possible to clearly distinguish protein activity at target *vs.* non-target plasmids using the assay. Numerous troubleshooting steps were taken to improve the assay conditions such as varying the amount of protein used, changing reaction buffer conditions and varying the reaction time, all with no success. Target plasmid comprising the *CDKN2B* promoter region was also evaluated in an attempt to mimic as closely as possible target site contact that these ZFs were ultimately designed to bind to. The promoter region of the *CDKN2B* gene was generated by PCR from the haematological cell line HL60 and subcloned into the pBL- CAT vector for the CAT reporter assay (This is described in more details in the next section (3.4) of this Chapter). The CAT plasmid harbouring the *CDKN2B* promoter (pBLCAT/*CDKN2B*) was used as a target vector and a pBLCAT empty vector was used as a non- target site vector control. These vectors gave the same result as previous LZF and LNS target vectors shown in Figure 3-10B.

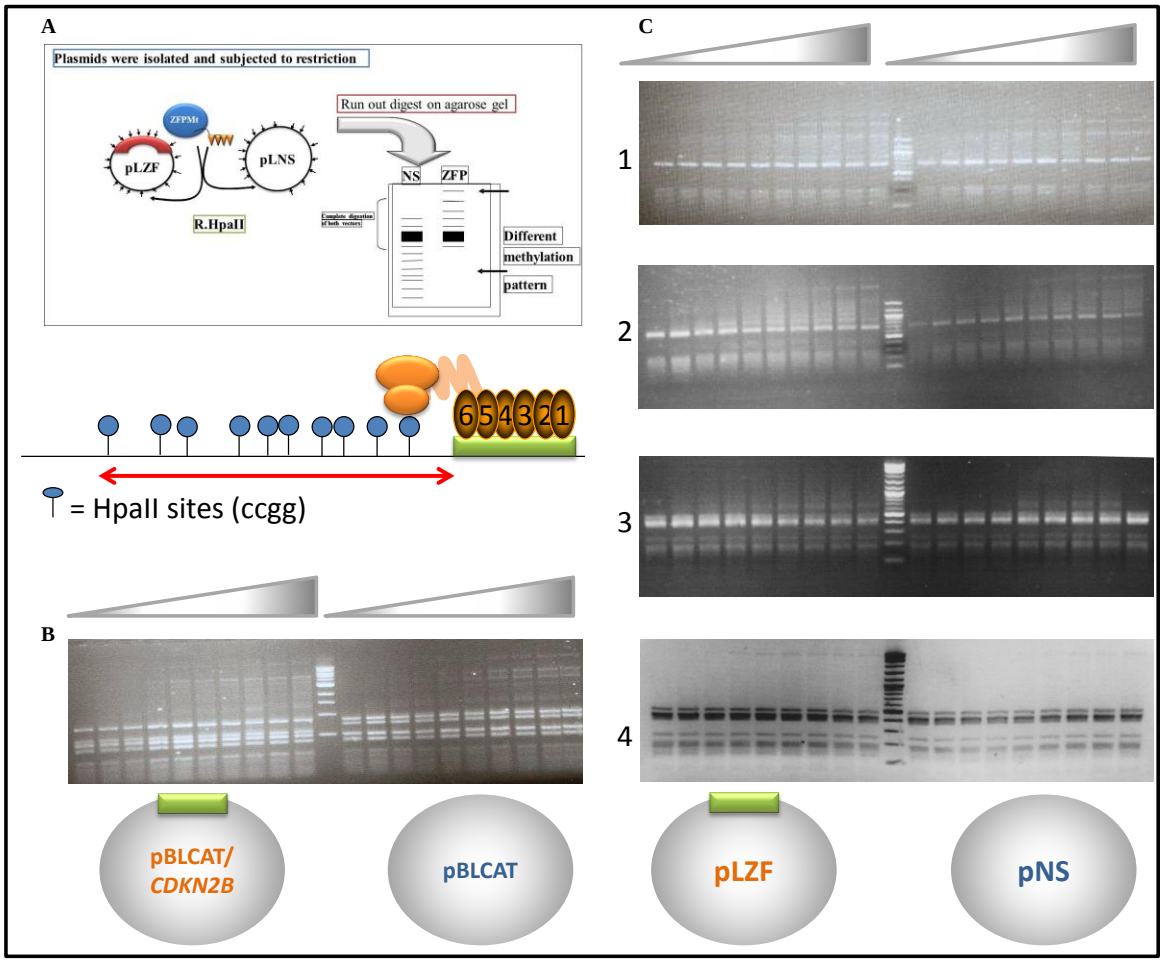


Figure 3-10. *In vitro* methyltransferase activity assay for 6ZFMIZ/HpaII, 6ZFAB/HpaII, 6ZFPU1/HpaII, and 6ZFAB/SssI proteins titrated with target plasmids pNS and pLZF pBL CAT/*CDKN2B* and pBLCAT.

A Diagrammatic representation of the basic principle of the *in vitro* methylation assay. **B** Titrated 6ZFAB/SssI protein was incubated for 45 minutes with the target plasmids pBL-CAT/*CDKN2B* and pBLCAT. This was followed by *R.HpaII* restriction to look for a unique restriction pattern consisting of higher molecular weight bands if the plasmid is methylated. Marker used was a 1kb DNA ladder. **C** 6ZFMIZ/HpaII (1), 6ZFAB/HpaII (2), 6ZFAB/SssI (3) and 6ZFPU1/HpaII (4) purified proteins were incubated for 45 minutes with the target plasmid pNS and pLZF. This was followed by heat inactivation of enzyme and *R.HpaII* restriction digest. Marker used was a 1kb DNA ladder.

3.4 CAT Reporter Activity Assay

The previous section described the attempts to assess the *in vitro* methyltransferase activity of the fusion proteins. I have demonstrated that the ZF/methyltransferase fusions we generated were capable of binding to their target sites in short oligonucleotides, but the information on the activity of the methyltransferase component was not conclusive. The next step was to target methylation using fusion proteins specific to the region of a *CDKN2B* regulatory region driving the Chloramphenicol Acetyltransferase (CAT) reporter gene in an *in vivo* setting. Following ZF/Mtase targeting, promoter transcriptional regulation was assessed by quantitatively measuring CAT expression in eukaryotic cells by CAT ELISA assay. It has previously been demonstrated that directing DNA methyltransferase activity to predetermined sites in DNA led to the transcriptional silencing/downregulation of *Renilla* luciferase reporter gene under control of the cytomegalovirus (CMV) promoter [295]. Although, the ultimate aim of this project is to target methylation to the promoter region of an endogenous gene, in the light of data from EMSA *vs*. *ex vivo* assays, reporter system evaluation was considered necessary.

CAT is a bacterial resistance gene that inactivates the chloramphenicol antibiotic by acetylating one or both of its hydroxyl groups. As eukaryotic cells do not possess this gene, it is an ideal candidate for reporter gene studies. Its activity is directly proportional to the rate of acetylation observed in the CAT assay. By monitoring CAT activity, it is possible to monitor accurately and sensitively transcriptional regulation of a promoter.

An alternative reporter system based on Green Fluorescent Protein (GFP) was also used driven by a high molecular weight *CDKN2B* promoter fragment. GFP is a 283 amino acid protein, which when excited, exhibits green fluorescence [296]. A fragment of the *CDKN2B* promoter (1206bp) was cloned upstream of the GFP coding sequence and output in terms of GFP expression was measured by FACS. In both settings, the expected outcome is that if the *CDKN2B* promoter becomes methylated, it should lead to the transcriptional downregulation of the CAT or GFP report system.

3.4.1 Cloning of *CDKN2B* promoter regulatory region

To investigate the targeted delivery of different CpG methylation patterns to gene regulatory regions *in vivo*, I constructed two reporter vectors driven by *CDKN2B* promoter regions. These PCR products contained all of the ZF binding sites as well as the most important regulatory sequences within the *CDKN2B* promoter, and the main difference between them is that the shorter 734bp PCR amplicon lacked the first 472bp of the 5' region shown in Figure 3-2 of this Chapter. Genomic DNA was isolated from the haematological cell line HL60 (Human promyelocytic leukaemia cells) as described in the Methods Chapter Section 2.8.3. The choice of the cell line was based on the fact that the methylation status of *CDKN2B* is known to be unmethylated in this case, unlike in the majority of haematological cell lines where this gene is either hypermethylated or deleted. Isolated genomic DNA was used as a template for PCR amplification to generate 734 bp and 1206 bp PCR fragments of *CDKN2B* promoter regions (Figure 3-11A).

The generated *CDKN2B* fragments were subcloned into pBLCAT or pGFP reporter system vectors for subsequent methylation targeting and gene expression analysis. In principle, the *CDKN2B* promoter targeted with ZFs lacking HpaII methyltransferase should have no effect on reporter gene expression. Conversely, targeting with ZFs containing methyltransferase component should lead to promoter hypermethylation mediated reporter gene silencing Figure 3-11B.

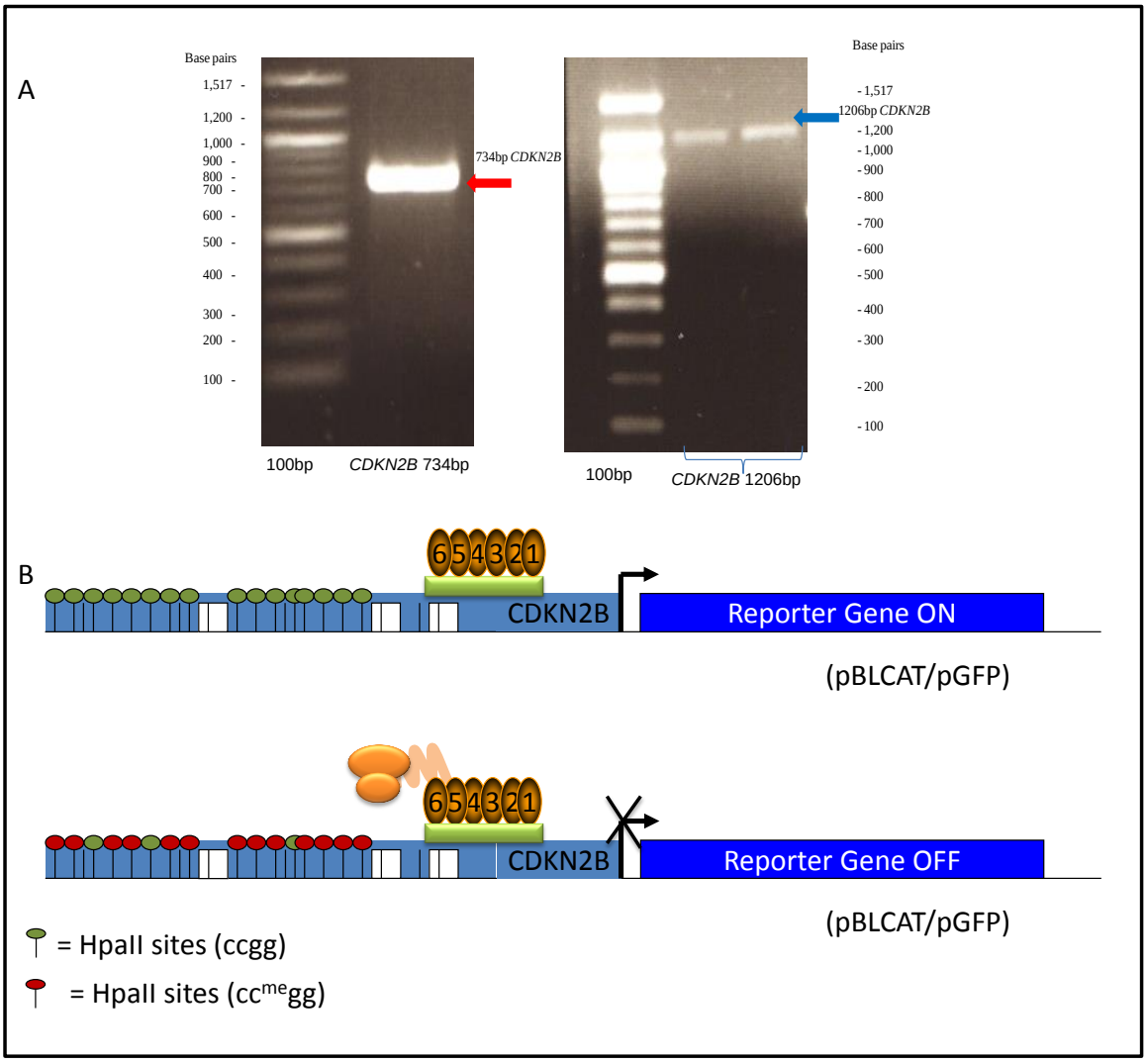


Figure 3-11. PCR Amplification of *CDKN2B* promoter regions

A *CDKN2B* 734bp fragment and **B** *CDKN2B* 1206bp fragment. Fragments were amplified from HL60 cells using primers with *XhoI/HindIII* overhangs. Following gel purification, PCR products were subcloned into pBLCAT and transfected into A549 cells. **B** Schematic diagram of basic principle behind the reporter assay showing *CDKN2B* promoter driving the reporter gene (GFP or CAT). Targeted methylation delivery to the *CDKN2B* promoter region should result in methylation associated gene repression.

PCR fragments were generated with XhoI/HindIII overhangs as described previously, for subsequent cloning into the chloramphenicol acetyltransferase (CAT) expression plasmid pBL-CAT. This resulted in two plasmids being generated, pBL-CATL12 (1206bp fragment) and pBL –CATS7 (734bp).

Zinc Finger: methyltransferase constructs were subcloned into a modified CMV FLAG mammalian cell expression vector as described in Chapter 4 Section 4.5, in preparation for endogenous *in vivo* targeting studies. ZF vectors used for this study were ACMVMZH, ACMVABH and ACMPU1H. Two cell lines; A549 (Human lung adenocarcinoma epithelial cell line) and HeLa (Human Cervical cancer) were used for transient co-transfection with pBL-CATL12 or pBL –CATS7 and one of the above described CMVZF vectors or non- specific 4ZFCMV/HpaII was used as a negative control. The A549 cell line was chosen for this purpose due to its demonstrated homozygous 9p21 segment deletion [297], which would exclude the possibility of background *CDKN2B* signal interfering with the assay readout. Due to the fact that A549 was difficult to transfect with high efficiency by nucleofection (using Amaxa pGFP plasmid, maximum efficiency was 20%), HeLa cells were selected as an alternative.

Transfected cells (A549 or HeLa) were harvested 96 h post-transfection, and CAT levels were determined using the CAT ELISA kit (Roche) using the protocol provided by the manufacturer. Briefly the cells were lysed in CAT ELISA kit lists buffer and cell extracts prepared. The basic principle of the assay is a sandwich ELISA. Lysed cells were added to the wells of the micro plate, already pre-coated with a polyclonal antibody to CAT (anti-CAT) (Figure 3.11). This anti-CAT antibody is bound to the plate surface and captures CAT protein present in the cell extract. Following the incubation step, the plates are washed 4x in CAT Elisa Wash Buffer and the digoxigenin-labeled antibody to CAT (anti-CAT-DIG) is added and allowed to incubate for 1 hour at 37°C. The plates were washed as previously described, followed by the addition of an antibody to digoxigenin, conjugated to peroxidase (anti-DIG-POD), which binds to digoxigenin.

Finally, the peroxidase substrate ABTS is added to catalyse the cleavage of the substrate, producing a coloured reaction sample. The absorbance of that coloured sample is determined using a spectrophotometer and is directly proportional to the level of CAT present in the cell extracts.

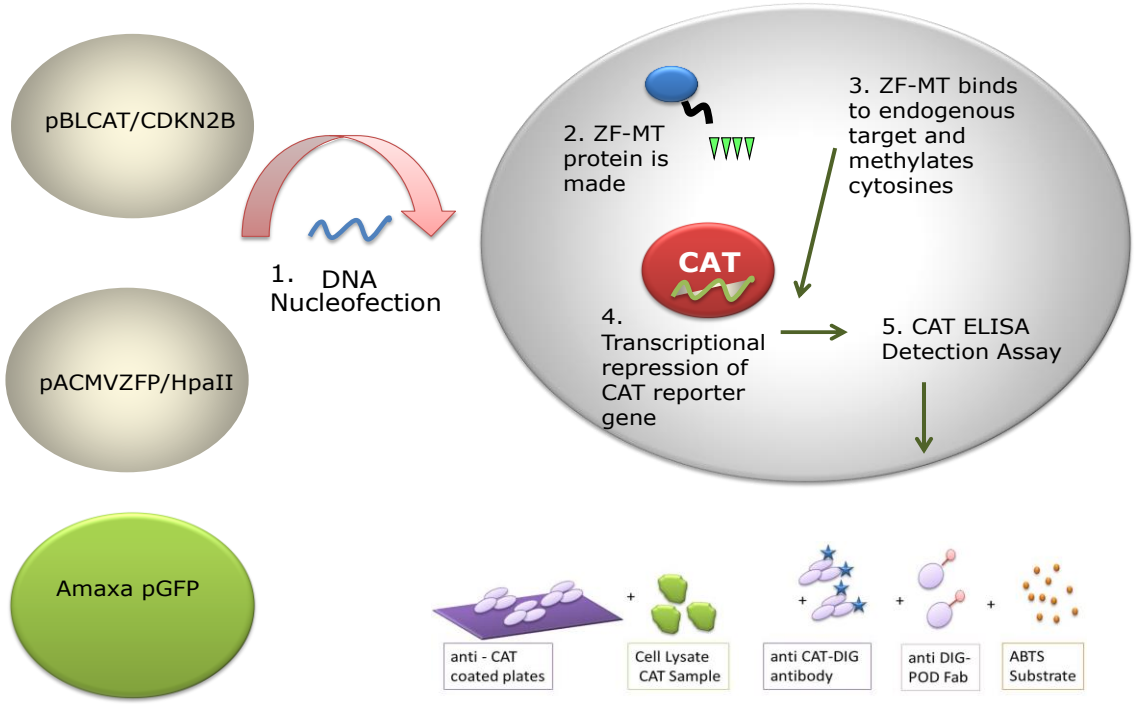


Figure 3-11. CAT Reporter Assay Description

Reporter vector pBLCAT vector, containing either a full or partial fragment of the *CDKN2B* promoter (1µg) was co-nucleofected with a CMV vector expressing one of the ZF/HpaII fusions (1µg) and an Amaxa GFP (0.5µg) plasmid to monitor transfection efficiency. Nucleofection was carried out as described in the Methods Chapter Section 2.10.3.1. Nucleofection Programmes were selected from the Amaxa cell database for 'High Efficiency'. The transfection efficiency was measured at 24, 48, and 72 h intervals. Four days following transfections, cells were evaluated by monitoring GFP fluorescence by microscopy and A549 efficiency was found to be 25% compared to 70% for HeLa cells. Therefore HeLa cells were selected for CAT Analysis using CAT ELISA.

Following the CAT ELISA assay no CAT expression could be detected. The maximum concentration of reporter plasmid available in nucleofection experiments was 2µg/transfection so I was limited when troubleshooting in changing the amount of plasmid used. An alternative approach was to change the transfection protocol and a number of different transfection procedures were investigated including Lipofectamine 2000™, Fugene™, Icafectin 441 ™, Jet PEI™ and CaBES as described in Methods Chapter 2 of this thesis. A CAT signal was eventually detected using CaBES and a Jet PEI™ method perhaps reflecting the fact that these methods allowed for the highest concentration of DNA (5µg) to be used (Figure 3.12). However, the results were not reproducible, and an extremely high amount of cell lysate protein was required to detect CAT signal by this method which made the assay impossible to execute.

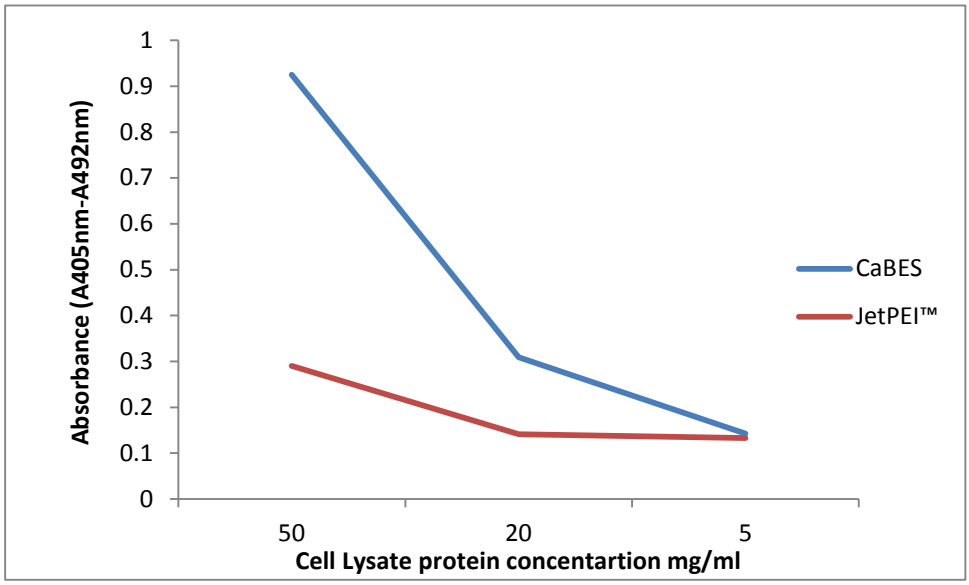


Figure 3-12. CAT ELISA curve following CaBES and JetPEI™ transfections.

Transfected HeLa cells ($1x10^6$) were lysed and the cell protein concentration determined by BCA. CaBES lysates were at 0.96mg/ml and the JetPEI™ cell lysate was at 3.95mg/ml.

Following unreliable detection of CAT protein expression, an alternative approach was to use the GFP reporter system. The *CDKN2B* (1206bp) fragment was subcloned upstream of GFP to generate the p15GFP vector. HeLa cells were co-transfected using the CaBES method under the same conditions as for the CAT assay. Cells were trypsinized four days following transfection and resuspended in DMEM media supplemented with 10% (v/v) FBS. GFP fluorescence was analysed by FACS (n=3). Unfortunately, only very low levels of GFP signal were detected (Figure 3.13).

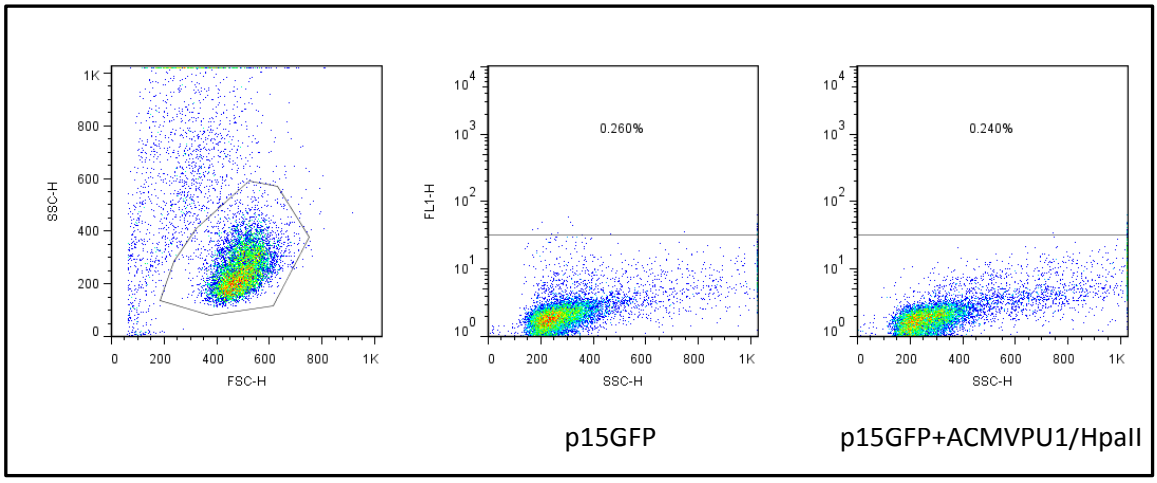


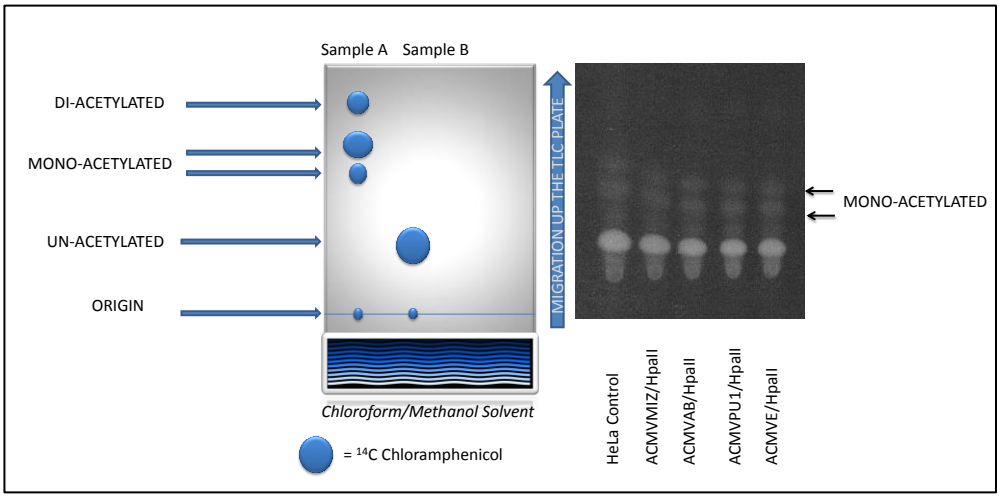
Figure 3-13. FACS Analysis of p15GFP transfected HeLa cells

HeLa cells ($1x10^6$) were transfected with 5µg of p15GFP only or 2.5µg of p15GFP and 2.5µgpACMVPU1/HpaII plasmid. 96h post transfection cells were collected for FACS analysis. GFP signal was gated on p15GFP transfected cells which were used as a negative control against a p15GFP+ACMPU1/HpaII transfected cells. Only a small increase in GFP signal of around 0.250% under both sets of conditions could be detected which was deemed insufficient for the reporter analysis following targeted methylation. FlowJo software was used to analyse transfection efficiency and the FL-1 channel was used to measure GFP intensity.

A final attempt at analysis of targeted methylation using an *in vivo* reporter system was made by changing the readout assay for the CAT reporter activity. Traditionally, CAT activity is measured using a radioisotopic CAT assay.

The radioactive CAT assay procedure was conducted as previously described. Briefly, a BCA assay was carried out to determine the concentration of the protein extract from the transfected cell lysate. Following incubation with ^{14}C chloramphenicol and co-factor Acetyl Coenzyme A, the complexes were spotted onto a thin layer chromatography plate (TLC) and placed in a chloroform/methanol solution. Following product migration through the TLC plates, when the solvent reached the top of the TLC sheet, the plates were dried and then placed into a Phosphorimager cassette overnight for quantitative imaging analysis. Imaging was performed using a Fuji Phosphorimager 400 that uses storage phosphor technology for visualization purposes [298]. Different states of chloramphenicol acetylation, which appear as spots of different mobility, appear in three forms – two forms of chloramphenicol with one acetyl group (monoacetylated), one form with two acetyl groups (di-acetylated) and all of the unused chloramphenicol (Figure 3.14). The density of the spots can then be scanned for quantitative analysis using AIDA 2D densitometry software. CAT assay was kindly conducted by Dr Ford.

As the result in Figure 3-14 shows, only un-acetylated forms of ^{14}C could be clearly detected. The band intensity of migrated forms was measured using AIDA densitometry and no significant differences were observed between various samples. This is in line with the findings of the CAT ELISA assay, as well as the findings using the GFP reporter system.



.

Figure 3-14. CAT Assay TLC schematic representation

Typical CAT assay result. The reaction is placed onto the origin spot, allowed to dry the TLC plate lowered into the solvent and allowed to resolve up the TLC sheet until the solvent reaches the top of the plate. The plate is then dried and the spots can be visualised using either X-ray film or Phosphorimager. Five samples prepared from transfections as described for the CAT assay were prepared run on the TLC plate. PBLCATL was used as a reporter vector.

3.5 Conclusion

Following successful subcloning and purification of different ZF/methyltransferase permutations, it was necessary to evaluate how fit for purpose newly generated fusion proteins were. The evaluation involved *in vitro* characterization of purified proteins for their binding affinity, specificity, *ex vivo* enzymatic activity, and ability to methylate and repress an artificial reporter system.

Titrative EMSA analysis employed to evaluate the binding affinity of fusion proteins showed efficient and specific binding to their target sites. Four ZF/Mtase fusions were evaluated and found to bind to their target site with varying degree of affinity and specificity to the ZF target site compared to non-specific targets. K_d values ranged from high for 6ZFPU1/HpaII, to exceptionally poor for 6ZFMYC/HpaII. The results for 6ZFPU1/HpaII and 6ZFAB/HpaII compare favourably to similar six finger protein evaluations previously conducted in our lab. Binding affinity data analysis showed that the PU1/HpaII binds to its target sites most strongly (apparent Kd 5.2nM); the AB/HpaII protein binds with intermediate strength (Kd 11nM). The Miz/HpaII fusion protein bound weak to its target site (Kd 43nM), and Myc/HpaII bound very weakly (Kd 992nM). Binding of the same fusions to the scrambled zinc finger and HpaII binding site was observed only at the extremely high concentrations of the proteins used (Figure 3-7). This indicates that the zinc finger dominates binding, while the non-target binding by the ZF: methyltransferase fusion protein a HpaII site only, occurs at very high concentrations of the proteins used. However, it should be pointed out that the Kd values given are a guide only, as variations in protein purity, concentration measurements, and probe concentration make such comparative measurements prone to error.

Given that most of the ZFs studied in these assays were designed and synthesized using the same approach, the differences in the binding affinities are perhaps surprising. Zinc finger proteins are modular in nature, with each module capable of triplet base subsite recognition. DNA:ZF interface interactions are believed to be highly diverse and complex, making attempts at developing a definite recognition code extremely difficult. In some cases, amino acid residues can recognize extended targets of four even five nucleotides [299]. These target site overlaps (TSO) can sometimes influence modularity of the adjacent fingers. For instance, in a modular

array with TSO issues one zinc finger can interfere with the target-site of the zinc finger adjacent to it. They are typically created by an aspartate residue at position 2 of the α-helix (observed in GNN, ANN, CNN and TNN recognition triplets) making contacts outside the triplet base [300]. This potential to specify overlapping 4bp subsites can affect inter-finger synergism as observed for Zif 268 where it was shown that deleting contacts from finger three can affect the adjacent binding site in finger two [301]. However, TSO analysis shown in ELISA graphs in Figure 3-3 A-C showed no TSO unmet preferences.

Following successful targeting of two of the ZFs to a specific DNA target gave a degree of confidence that once fused with various methyltransferases for in vivo endogenous gene targeting, these proteins would be able to recognize a contiguous 18bp sequence within the mammalian genome.

However, *in vitro* enzymatic activity evaluation was disappointing as all attempts to target methylation to the HpaII sites present on a ZF site containing plasmid did not succeed. Previous results in our lab, when using these assays, have shown a clear trade-off between the binding affinity of the ZF component of the fusion protein and the methyltransferase enzyme activity. It may be that the binding of ZF with high affinity could compromise/restrict methyltransferase activity leading to the observed non - targeted methylation seen in these assays. This statement is based on the observed requirement for a large conformational change to occur in the methyltransferase component, in order for the methyl transfer reaction to occur.

Another important point to consider is the orientation of the methyltransferase target in relation to the ZF binding site. This can often determine the success of targeting of the ZF/methyltransferase protein, especially if the HpaII sites are some distance away from the ZF site. The topology of the target vectors also needs to be considered as a contributing factor to the observed results. It is possible that supercoiled plasmid does not allow the ZF to 'see' its target site, or alternatively might present a high affinity substrate for the Mtase component, overriding the target properties of the ZF component.

Final point to consider is the flexibility between the fused domains of this enzyme. In this study, fusion proteins were joined by the flexible linker, $(Gly_4Ser)_3$. It has previously been reported that the use of linkers in fusion proteins can affect their folding stability due to issues of linker flexibility and hydrophilicity. The linker used for this project has been demonstrated to have an optimum

length and composition [302]. This is because Glycine has no unfavourable β – carbon allowing the polypeptide backbone to access dihedral angles which are energetically forbidden to other amino acids [303]. Therefore, a linker which is glycine rich is more flexible than a linker of the same length composed of other residues. Conversely, too much linker flexibility will affect the single chain protein stability, which was avoided by inclusion of Serine residues in the linker. Studies using serine containing linkers indicate that there is a decrease in the unfolding rate of the protein inversely proportional to the increase in the number of the serine residues used. Serine residues within the linker will also increase hydrophilicity due to the ability of serine to form hydrogen bonds, allowing the formation of new stabilizing interactions [302].

Unfortunately, attempts to characterize the ability of ZF/Mtase fusions to target methylation to the *CDKN2B* promoter region of the CAT and GFP reporter constructs *in vivo* were not successful. The *CDKN2B* fragment was amplified from HL60 cells, and the region selected was believed to contain all the key regulatory sequences. Unfortunately, very low to undetectable levels of reporter activity (CAT or GFP) in HeLa cells or 293 cells were detected during this study. I have found an example of a study using *CDKN2B* promoter reporter system. In this study, various *CDKN2B* fragments were used to evaluate Smad 4 recruitment to the promoter region in Capan-1 and CFPac-1 human pancreatic cancer cells. Two of the fragments they studied covered the region evaluated in my own study and showed a significant luciferase reporter activity following Smad4 binding to *CDKN2B*. However, the luciferase construct they used, p3TP-lux also contained a 3TP promoter with three consecutive TPA response elements (TREs) and a segment of plasminogen activator inhibitor 1 (PAI-1) [304]. An alternative study on Sp1 mediated activation of *CDKN2B* used a *CDKN2B* driven luc reporter system successfully only upon the TGFβ activation [305].

In conclusion, the promoter fragments used in this thesis were designed to cover the region between - 863 to + 273 region of *CDKN2B* with the long fragment 1206bp and a -383 to + 273 for the short fragment of the *CDKN2B*. The *CDKN2B* promoter linked to the reporter system CAT or GFP used in our study was either too weak or ineffective in generating sufficient signals for further analysis.

Chapter 4

Targeting DNA methylation to the endogenous *CDKN2B* gene

4.1 General Introduction

Constitutively expressed and some tissue and lineage specific genes have CpG islands (CGIs) co-localized with their promoters, leading in part to the observed uneven distribution of CpGs within the genome [306]. These CG rich regions can harbour the transcription start site (TSS) of genes and are usually associated with transcriptionally permissive chromatin [307].

There are suggestions that CGIs serve as a recruitment locus for transcription factor (TF) binding. This is based on an analysis of the density of transcriptional elements in the promoter and non-promoter CGI sequences, which showed around 40% relatively higher enrichment for TFs in promoter regions [308].

Characterization of the chromatin structure associated with unmethylated CGIs reveals a strong correlation between CpG density and the presence of specific histone marks of actively transcribing genes *i.e.* H3K4 trimethylation, and H3 acetylation. Further support for a mechanistic interdependence between chromatin modifications and CGIs comes from observations that certain CpG binding proteins such as Cfp1, which exclusively associate with non-methylated CpGs, also interact with the H3K4 methyltransferase SETD1, to drive establishment of H3K4me3 [309].

A majority of CGIs can be found within the 5'-untranslated region or exon 1 of genes, with only a few exceptions found in the gene body or 3'-region. It has been suggested that these atypically located CGIs are more likely to be DNA methylated [310]. In the majority of cases, promoter based CGIs are maintained methylation free, correlating with gene expression.

Methylation of CpG rich promoters is often associated with down regulation of gene expression. A common question is whether such observed methylation is a cause or consequence of gene silencing. One proposed mechanism of gene silencing by methylation is through reduction of binding affinity of transcriptional activator for its target site. Site-specific methylation may block access to DNA of TFs required for optimal expression. For instance, adenoviral TF Major Late Transcription Factor (MLTF) is prevented from binding its target site as a consequence of

methylation [311]. Another possibility for methylation mediated gene silencing is through recruitment of methyl CpG specific DNA-binding proteins which bind to methylated CGIs *in vivo*. It is assumed that this blocks access to the transcription factors necessary for transcription initiation. This model also implicates histone deacetylase activity in methylation mediated transcriptional repression, since this activity is shown to be recruited by methylated-DNA-binding proteins, such as MeCp2 [312, 313].

CGI hypermethylation of a tumour suppressor gene in a human cancer was first described for the Retinoblastoma gene (*RBI)* [314] and several years later, it was proposed as a mechanism of gene inactivation of the Von Hippel-Lindau (*VHL*) gene [315]. Hypermethylation of the CGI of the *CDKN2A* gene, which encodes an important cell cycle regulator ($p16^{INK4A}$) occurs frequently in human cancers and is associated with gene silencing [316-318]. The identification of a large number of hypermethylated genes such as *CDKN2A*, *CDKN2B*, *TP73*, *BRCA1*, *hMLH1*, *GST3*, *MGMT*, *CDH1*, *TIMP3*, and *DAPK* followed, in a study of over 600 large primary tumour samples, representing 15 major tumour types [319]. Another large study of sporadic and inherited breast and colon tumours showed that the cancer cell genome undergoes a reduction of its 5-methylcytosine content compared with normal tissue of the same origin [320]. This finding is difficult to explain considering that there is an overall increase in the activity of the DNMTs (DNMT1, DNMT3A and DNMT3B), which are shown to be upregulated in solid tumours and haematological malignancies such as acute and chronic myeloid leukaemia [321, 322].

One explanation for this paradoxical distribution of DNA methylation throughout the genome might be that DNMTs somehow become unable to recognize sequences such as DNA repeat regions or intronic sequences that are normally required to be methylated, whereas CGIs, which are not usually targets of the DNMTs in a chromatin context, become methylated. In this regard, it is perhaps noteworthy than mutations in many cellular epigenetic factors (*e.g. Bmi*I, *Ezh2*, *Asxl1*), including DNMTs, are now being observed in cancer and leukaemia patients, and these may be informative of the mechanisms behind aberrant DNA methylation in these diseases. Apart from imprinted genes [323] and X-chromosome genes in females [324], the majority of CGIs are unmethylated in normal tissue.

As it stands, the question of what triggers the decision to include or exclude methylation in the cell, and how DNA methylation patterns arise in the cancer cell, specifically the hypermethylation of normally unmethylated gene promoter regions, remains unanswered. Studies described below suggest that CpG methylation of a promoter may be a progressive process, where several cycles of aberrant methylation may need to occur before a dense hypermethylated pattern can be imposed, leading to transcriptional silencing. DNA methyltransferase enzymes primarily target CpG dinucleotides, but paradoxically even during development where global *de novo* methylation changes occur within the epigenome, CGIs in healthy cells are usually unmethylated [325]. Two mechanisms have been proposed to explain the occurrence of the aberrant *de novo* methylation in cancer. The first mechanism proposes methylation spreading into the CGI from a *de novo* methylation centre juxtaposed to that island, such as the flanking Alu repeats surrounding the *CDH1* and *VHL* tumour suppressor genes [326]. The second hypothesis is that there are methylation hot spots within the CGIs which succumb to methylation, leading to the spread until a state of hypermethylation is achieved. This has been observed for the glutathione-S-transferase P1 (*GSTP1*) gene promoter in prostate cancer, where a combination of random site methylation and prior gene silencing lead to promoter hypermethylation [327]. Both mechanisms do not exclude the possibility that CGIs are kept methylation-free by the existence of certain boundaries which provide protection from methylation, perhaps by binding interference and which are progressively overridden during the disease process.

Studies on the mouse and hamster adenine phosphoribosyltransferase (*APRT*) housekeeping gene promoter for example showed that the island is free of methylation, with the flanking regions containing methylated CpG sites. This study proposed a model by which *APRT* is kept methylation- free, most likely *via* Sp1 occupying 5' sites of the promoter [328, 329]. It is also possible that RNA transcripts anti sense to a promoter may protect the CGIs either directly or in combination with 5MeC-DNA glycosylase by demethylation targeting [330]. Finally, there is a suggestion that specialized nucleosomal structures, chromatin signatures and/or replication timing may inhibit CGI methylation [331]. CGIs are believed to be sites of both transcription and replication, based on their observed co-localization with the origin of replication [332]. Mammalian replication origins that co-localize with CGIs were observed for a number of genes such as Heat shock protein 70, *c-MYC*, and *CDKN2B* [333-336]. This observation has led to the

suggestion that replication initiation intermediates can lead to the exclusion of DNA methylation [337]. Issues of promoter methylation and protection will be discussed in more detail in the next chapter.

It appears that some CGIs become hypermethylated, conferring a selective advantage for the survival of the cancer cell. An example for this Darwinian concept is the *BRCA1* gene found to be hypermethylated only in breast and ovarian cancer [338], or the mismatch repair gene *hMLH1*, which is epigenetically silenced only in colon, gastric and endometrial neoplasia where its transcriptional downregulation has important cellular consequences [339]. These findings correlate very well with the familial tumour genetic studies on *BRCA1* and *hMLH1* germline mutations. However, the situation is not that simple, which is exemplified by two illustrative cases; 1) in breast cancer only *BRCA1* gets hypermethylated while *BRCA2* does not [340]; 2) the *hMLH1* promoter is hypermethylated, while closely related family members with very similar CGIs such as *hMSH2*, *hMSH6* and *hMSH3* are not [339]. It is not clear whether a specific property of various CGIs protects them from methylation or if there is a more general shield that prevents access to DNMTs. The correlation between gene silencing and promoter hypermethylation is not valid in all cases. Gene silencing can occur prior to any *de novo* DNA methylation or chromatin remodelling event. This was demonstrated for the *CDKN2A* CGI in primary human mammary epithelial cell (HMEC) lines during selection, where gene silencing was shown to precede epigenetic suppression and low levels of *de novo* methylation were associated with a dynamic remodelling of histone modifications [341].

In order to investigate the role of DNA methylation signatures at gene promoter regions I needed to reproduce methylation patterns at specific CGIs. Previously it was shown in our lab that it is possible to target DNA methylation to an artificial CGI and it was confirmed that such imposed methylation was able to down regulate gene expression and in a heritable manner [342]. This method of targeting DNA methylation involved expressing zinc finger proteins expressed as fusions with the catalytic domains of mammalian or prokaryotic methyltransferases [343-347].

4.2 Zinc Finger Protein Methodology

The modular nature of individual zinc fingers within ZFP arrays has suggested ZFPs as appropriate templates for the construction of highly specific gene targeted binding proteins. The generation of ZFP:DNA recognition codes (explained in detail in Chapter 3) has enabled the design and construction of zinc fingers designed to bind virtually any DNA sequence [348]. By stringing zinc finger modules in a tandem array for example, unique sequences of up to 18 base pairs have been constructed for targeted binding *in vivo* [349]. ZFPs with such engineered DNA sequence specificities have so far been used either without any functional domain attachments or as fusions with various effector domains that directly interact with DNA and/or with other proteins. These include ZFPs for gene targeting and regulation purposes tethered to cleavage domains of endonuclease FokI [350], functional domains of transcription factors such as Sp1 [351], chimeric transposases [352], chimeric recombinases with the catalytic domain from bacterial Tn3 resolvase [353], HIV retroviral integrase [354], repressor domains such as Kruppel associated box (KRAB), ERF repressor domain (ERD), TATA box binding protein (TBP), transcription activation domain of herpes simplex virus VP16 region [355-359], or the p65 domain of the human endogenous transcription factor NFκB [360] and various methyltransferases [361].

Some of the above mentioned examples are of ZFP-based transcription factors targeted to artificially integrated chromosomal loci, or to non-chromatin targets such as targeting of HSV-1 LTR using six finger arrays [362, 363]. There are however also clear cases where artificial ZFP-based TFs have been shown to regulate endogenous gene expression. These target genes include human erythropoietin (*EPO1*) [364], Vascular Endothelial Growth Factor A (*VEGF-A*) [365], the erbB-3 protoncogene [366], while the peroxisome proliferator activated receptor-γ (*PPARγ*) [367], multidrug resistance gene (*MDR1*) [358] and checkpoint kinase 2 (*CHK2*) [368] have all been demonstrably downregulated by artificial ZFP-based repressors. Some of the ZFP fusions such as the ZFP transcription activator VEGF-A fusion (SB-509), ZFP fusions with CCR5 co-receptor which is used by the HIV virus to infect the cells of the immune system (SB-728), and ZFP/ glucocorticoid receptor gene fusion (SB-313) have been through Phase I and II clinical

trials which have proved their safety and efficacy in the treatment of diabetic neuropathy [369, 370].

The first attempt to target a methyltransferase component fused to a zinc finger (ZF) was to the p53 DNA binding site of the *CDKN1A* gene. This method was shown to be capable of delivering methylation to the desired target regions in a site-specific manner in *in vitro* assays [371]. Similar targeting effects were observed in a yeast *in vivo* model examining the *CAR1* promoter, using ZFP/M.SssI or M.CviPI fusions [372]. Additionally, it has previously been demonstrated in our lab that three ZF, M.HhaI or M.HpaII fusion proteins resulted in specific DNA methylation *in vitro* [346]. Extension of this work to the genomic context using cell lines harbouring integrated reporter genes driven by artificial CG rich minimal promoter regions confirmed targeted methylation specificity, gene repression due to targeted DNA methylation and the heritability of the silencing signal [342]. This project aims to take this process one step further by delivering targeted methylation to an endogenous gene promoter relevant to a specific disease. The success of intracellular targeting by ZFP-Methyltransferases (ZFP-MT) is highly dependent on the correct delivery of the zinc finger in its properly folded form and the ability of that finger to access its intended target.

4.3 *CDKN2B* Context

The *CDKN2B* gene represents a challenging and important model candidate gene for targeting studies, based on previous observations that this is one of the most frequently methylated genes in particular subtypes of haematological malignancies, such as MDS, AML and ALL. *CDKN2B* encodes $p15^{INK4b}$, a cyclin-dependent kinase inhibitor that is expressed selectively in haematopoietic progenitor cells during myeloid and megakaryocytic, but not erythroid, differentiation [373, 374]. Aberrant promoter hypermethylation occurs in the promoter region of the *CDKN2B* gene and is associated with its transcriptional silencing [375-377].

The *CDKN2B* CGI has a 67% GC content and structurally is a good candidate for attempts at methylation-mediated silencing due to the fact that the transcriptional start site is located within

the island. This region contains a number of regulatory element binding sites as well as methylation sensitive restriction endonuclease sites, such as seven HpaII sites (with recognition sequence CCGG) and sixteen HhaI sites (GCGC), which historically have been used to probe the methylation state of the *CDKN2B* promoter [378, 379].

4.4 Chapter Aim

I have rationally designed six zinc finger proteins targeting different regions of the promoter of the endogenous *CDKN2B* gene. The GCG Wisconscin Package TF software was used to determine the putative location of important regulatory sites within the promoter in the region spanning ~1500 base pairs. Also, target sites were identified from published work relating to the key transcriptional regulator binding sited for *CDKN2B*. Precise ZF sites were chosen after considering their proximity to these regions and the transcription start site. Zinc finger protein ZMZ, was designed to recognise the 5' flanking region of the promoter at +238 from the Initiator Sequence (Ir), which is relatively sparsely populated with CG residues (for details see Figure 3-2 Chapter 3). This region harbours a Miz-1 binding site, a transcription factor putatively involved in up regulation of *CDKN2B* [380]. The rationale for targeting this relatively CG poor region of the promoter was based on an earlier observation that in certain promoters, such as that of the *Interleukin 2* (*Il2*) gene, which is silent in naïve murine T lymphocytes but becomes expressed upon T cell activation, only the most proximal CG residues are proposed to play a role in methylation related gene regulation [381]. Additionally, this region of the *CDKN2B* promoter also contains putative *Smad* binding sites, which are transcription factors directly activated by transforming growth factor-β (*TGF-β*) and involved in *CDKN2B* transactivation via *Miz-1* [382].
Two additional six zinc finger proteins, ZMC and ZAB, were designed to recognise sites within the most densely CG populated region at -186 and -197 (Figure 1.20 Chapter 1). This region contains binding sites for many important regulatory elements, such as *Sp-1*, a transcriptional activator of the *CDKN2B* gene [383]. Methylation profile analysis of the *CDKN2B* gene in samples from patients with AML, ALL, and CML shows frequent methylation at the aforementioned *Sp-1* site [384]. Although *Sp-1* itself is not sensitive to DNA methylation, binding

to DNA could be blocked by recruitment of methyl-cytosine binding proteins, as discussed previously in Chapter 1.

A fourth six zinc finger protein ZPU1 was designed to recognise a sequence further downstream at -402 in the 3' promoter region close to the *CDKN2B* translation start site. Following rational design, all of the constructs encoding ZFPs were synthesized commercially by GenScript, with the exception of ZFPAB, which was assembled using the SOE-PCR approach previously described in Chapter 3.

4.5 CMV FLAG-TAG Expression

Zinc Finger Proteins targeting the *CDKN2B* promoter region were rationally designed using the Barbas' lab Zinc Finger Tools website as specified previously in Chapter 3 Section 3-2. [385]. There are 64 DNA triplets that could be potentially targeted by ZFPs and up to 10kb of DNA sequence can be searched at once to select potential, high affinity ZFP binding sites. The amino acid sequence of each zinc finger protein used in these studies as well as the rationale for the ZFP selection based on the scoring system was described in Chapter 3 Section 3.2. and Figure 3-6.

The commercially synthesized and sequenced ZFPs were evaluated in *in vitro* settings for their binding affinity. ZFPs were cloned into human cytomegalovirus (CMV) based vectors and expressed either alone or as amino terminal fusions with the prokaryotic methyltransferases M.SssI, M.HpaII and with the catalytic domains of the murine DNMT3A and DNMT3B methyltransferase, for the endogenous targeting studies. CMV expression vectors contained the nuclear localization signal (NLS) PKKKRKV and an N-terminal FLAG tag, for the detection of protein expression. The function of the NLS signal is to direct the protein to the cell nucleus through the nuclear pore complex *via* its recognition by cytosolic nuclear transport receptors. The general schematic for ZF-MTase constructs is shown in Figure 4-1. The list of all of the constructs designed this way is shown in Table 4-1.

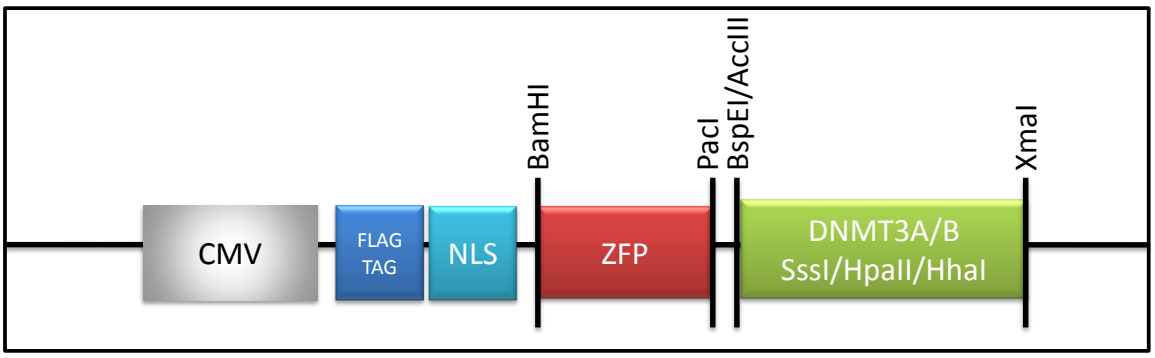


Figure 4-1. Schematic of expression constructs for ZF-MTases.

The schematic diagram describes the basis of Flag-NLS-ZF-Mtase expression constructs used for the studies. Various Mtases are fused with ZFPs, a nuclear localization sequence (NLS) and a FLAG-tag. All ZFPs were synthesized to contain *BamHI* /*PacI* overhangs to enable subcloning in frame with different methyltransferases. DNMT3A and DNMT3B methyltransferase catalytic domains were subcloned as *BspEI*/*XmaI* fragments, whereas prokaryotic M.SssI and M.HpaII were subcloned as *AccIII*/*XmaI* fragments. The constructs were sequence verified (Cogenics).

ZPU1	ZMZ	ZAB	ZMC
ACMVZPU1B	ACMVZMZB	ACMVZABB	ACMVZMCB
ACMVZPU1S	ACMVZMZS	ACMVZABS	ACMVZMCS
ACMVZPU1H	ACMVZMZH	ACMVZABH	ACMVZMCH
ACMVZPU1A	ACMVZMZA	ACMVZABA	ACMVZMCA
ACMVZPU1Hh	ACMVZMZHh	ACMVZABHh	ACMVZMCHh
ACMVZPU1	ACMVZMZ	ACMVZAB	ACMVZMC

Table 4-1. CMV constructs of ZFP/methyltransferase fusions.

Four ZFPs (PU1, MIZ, AB and MYC) were fused with either eukaryotic methyltransferase catalytic domains DNMT3B and DNMT3A or prokaryotic methyltransferases M.SssI, M.HhaI and M.HpaII in a CMV vector for mammalian expression and targeting studies. CMV constructs containing only ZFPs without methyltransferases were also generated for use as a negative control in targeting studies.

4.6 CMV FLAG Detection

Initial experiments were designed to assess whether HeLa cells transfected with the constructs described above expressed FLAG tagged protein fusions of the correct size. The FLAG–tag is a polypeptide tag with the sequence N-DYKDDDDK-C, 1012 Da in size which is detectable with an anti-FLAG monoclonal antibody [386]. The tag has a wide variety of applications, primarily in affinity chromatography as well as recombinant protein isolation and purification and is not known to be expressed as a part of any human gene. FLAG-TAG was originally developed as a small N-terminus hydrophilic peptide engineered for the purpose of detection and purification of recombinant lymphokines from yeast supernatants or *E. coli* extracts [387]. Constructs listed in Table 4-1 were maxi prepped as described in Methods Chapter 2 Section 2.8.9.2. and transfected into HeLa cells, using the CaBES method described previously in Methods Chapter 2 Section 2.11.3.3. Samples were taken four days post transfection for western blotting. Transfected HeLa cell samples at $1x10^6$ were trypsinized, washed in PBS, and western blotted as described in Methods Chapter 2 Section 2.13.2. Novex ® Sharp Protein Standard was used

for molecular weight estimation. The marker covers 3.5-260kDa range. Monoclonal ANTI-FLAG® M2 Clone antibody (Cat No F3165, Sigma-Aldrich) was diluted 1/2000 in PBS with 1%BSA (w/v).

The expression from each of the constructs shown in Figure 4-1 in HeLa cells is shown in Figure 4-2 and the cellular localization of the ZFPs/ methyltransferase fusions by immunofluorescence in Figure 4-5. The FLAG-tag attached to the protein is highly hydrophilic and as such less likely to denature or inactivate the proteins it is attached to [388].

FLAG-tag expression was confirmed by western blot and the results are shown in Figure 4-2. Expression of FLAG-tag ACMV constructs containing ZF alone (ACMVPU1, ACMVMC, ACMVMZ and ACMVAB) was confirmed with the observation of an expected band of 21.8 kDa, Panel **A**. Also shown in Figure 4-2 **A** is confirmed expression of zinc finger protein AB fused with either DNMT3A or DNMT3B catalytic domain methyltransferase with expected band sizes for ACMVABB 56.3 kDa and 58.5 kDa for ACMVABA. Figure 4- 2 **B** shows three of the ZFPs (PU1, MYC and MIZ) either alone or as a part of DNMT3B catalytic domain fusions. The expression of FLAG-TAG ACMVPU1B, ACMVMZB and ACMVMCB ZFP/methyltransferase fusions was observed with the detection of a 56.3 kDa band size and a 21.8 kDa band size corresponding to ACMVPU1, ACMVMZ and ACMVMC zinc fingers alone. Finally, the expression of the same constructs with the DNMT3A catalytic domain methyltransferase was confirmed with the observation of 58.5 kDa band for ACMVPU1A, ACMVMZA and ACMVMCA.

Vector Used	Protein	Protein Size Expected (kDa)
ACMVPU1	FLAG-PU1	21.8
ACMVPU1B	FLAG-PU1B	56.3
ACMVPU1A	FLAG-PU1A	58.5
ACMVPU1H	FLAG-PU1H	62.0
ACMVPU1Hh	FLAG-PU1Hh	58.8
ACMVPU1S	FLAG-PU1S	63.8
ACMVAB	FLAG-AB	21.8
ACMVABB	FLAG-ABB	56.3
ACMVABA	FLAG-ABA	58.5
ACMVMZ	FLAG-MZ	21.8
ACMVMZB	FLAG-MZB	56.3
ACMVMZA	FLAG-MZA	58.5
ACMVMC	FLAG-MYC	21.8
ACMVMCB	FLAG-MCB	56.3
ACMVMCA	FLAG-MCA	58.5

Table 4-2. FLAG-TAG ZFP/methyltransferase fusion proteins.

This table summarizes nomenclature of FLAG- ZF/ methyltransferase fusion proteins and their respective sizes. Proteins were detected by western blot.

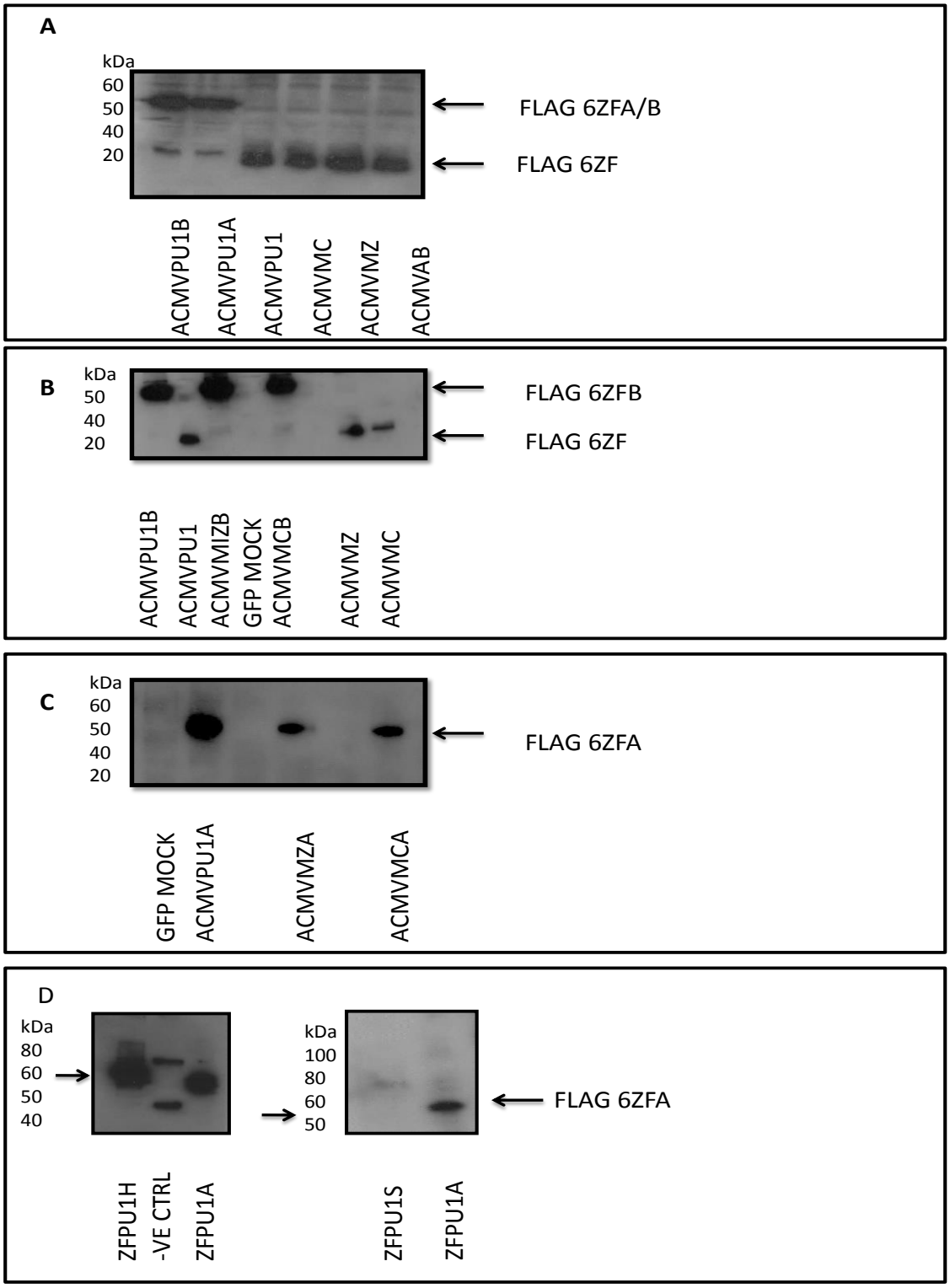


Figure 4-2 FLAG- tag protein expression in HeLa cells.

FLAG- ZFP/methyltransferase permutations proteins were detected using anti FLAG-tag monoclonal antibody following transfection of HeLa cells. The constructs shown in Table 4-1 were used to transfect HeLa cells and four days post transfections samples were collected and western blotted for the detection of FLAG-tag expression.

Although a weak band can be seen for ZFPU1S in Figure 4-2, no consistent FLAG- tag signal could be detected in subsequent experiments. The methylation status of these plasmids was checked using the methylation sensitive restriction enzyme *R. HpaII* (CCGG) and its isoschizomer *R. MspI*, which is not sensitive to methylation (Figure 4-3). More than 50% of the plasmid is methylated, i.e. remains uncut using HpaII methylation sensitive restriction enzyme. It is not unlikely that significant levels of plasmid methylation may lead to lack of expression from the plasmid itself in transient assays. The plasmid also has more potentially methylatable CpG sites compared to HpaII and HhaI vectors.

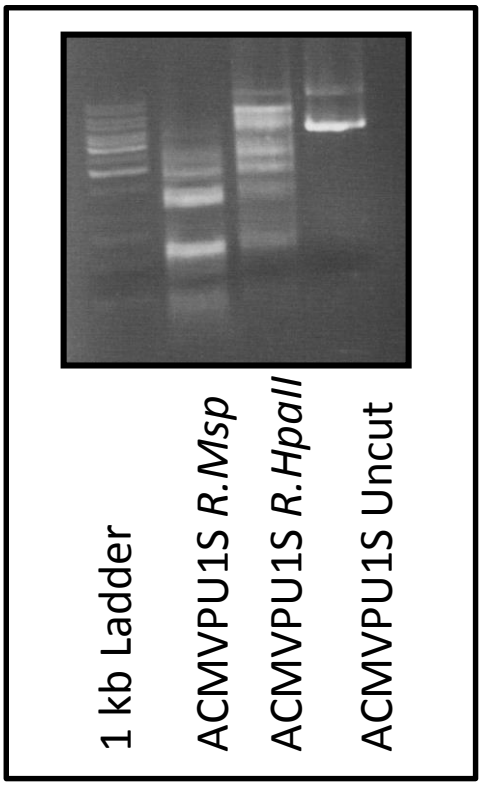


Figure 4 -3 Restriction digest analysis of CMVZFPU1S construct.

CMV ZFP/SssI plasmid construct digested with methylation sensitive *R.HpaII* and insensitive *R.MspI* enzymes. Cleavage by *R.HpaII* is obstructed indicating that the plasmid has become methylated.

To better understand and assess the adaptation of the bacterial genes (M.HpaII, M.HhaI, M.SssI) to their mammalian hosts (HeLa cell line) as well as to possibly predict the levels of their expression and gain an insight into the likely success of heterologous gene expression a Codon Adaptation Index (CAI) was used [389]. Comparisons of codon usage, GC content and the percentage of low frequency (<30%) codons of the methyltransferases used for this projects; M.HpaII, M.HhaI, M.SssI, and DNMT3B (Catalytic domain) were assessed using GenScript Rare Codon Analysis Tool (http://www.genscript.com/cgi-bin/tools/rare_codon_analysis).

Codons can influence the translation rate of a protein in a number of ways. CAI is a useful tool for predicting the level of gene expression based on the pattern of codon usage [390]. It has been observed in *Escherichia coli*, that if a target protein is not derived from the same species as the expression host its codon usage differs significantly from that of the host organism. Using two different DNA fragments, one with common codons and the other with infrequent codons inserted into the lacZ gene, a six fold difference in translation rate between the two has been observed [391].

A CAI value of 1.0 is considered ideal with any significant deviations from this value representing a potential negative effect on gene expression. The GC content of the gene can have an influence on transcriptional efficiency of that gene and the efficiency of translation of its mRNA [392]. Generally, GC content ranging between 30% and 70% is considered best. Whilst the favourable codons are read by the most abundant tRNAs and therefore highly translated [393], tandem rare codons are read by tRNA that are present in low abundance [394], which can lead to stalling of translation and premature degradation of mRNA [395, 396]. Figure 4-4 summarizes codon usage information for methyltransferases used in this project.

The CAI, GC content and codon usage by DNMT3B catalytic domain eukaryotic methyltransferase was used as a reference point against the prokaryotic methyltransferases (M.HpaII, M. SssI and M.HhaI). The only significant difference was in GC % content between the DNMT3B (50%) and the prokaryotic methyltransferase M.SssI (27%), M.HpaII (34%) and M.HhaI (33%).

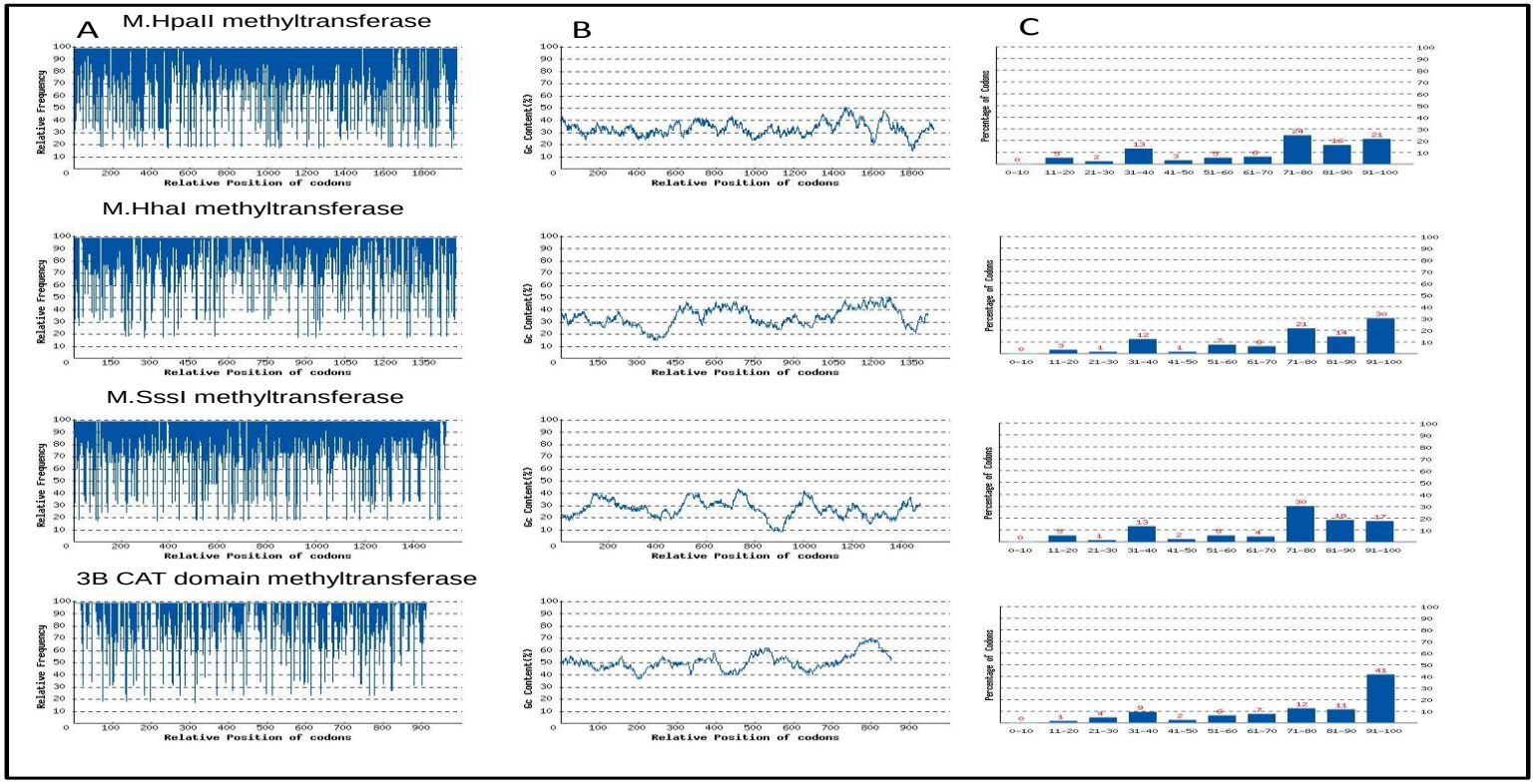


Figure 4-4 Codon Analysis Information

Graph representation of the CAI values, GC values and Codon Usage for the three prokaryotic methyltransferases used M.HpaII, M.HhaI and M .SssI and a catalytic domain of DNMT3B eukaryotic methyltransferase. CAI values **A**, CG content **B** and codon usage **C** for the above methyltransferases is listed in Table 4-3 below.

Methyltransferase	CAI	CG%	Codon Usage%
M. SssI	0.63	26.91	6
M. HpaII	0.68	33.64	4
M. HhaI	0.63	32.79	7
CAT DNMT3B	0.70	50.48	5

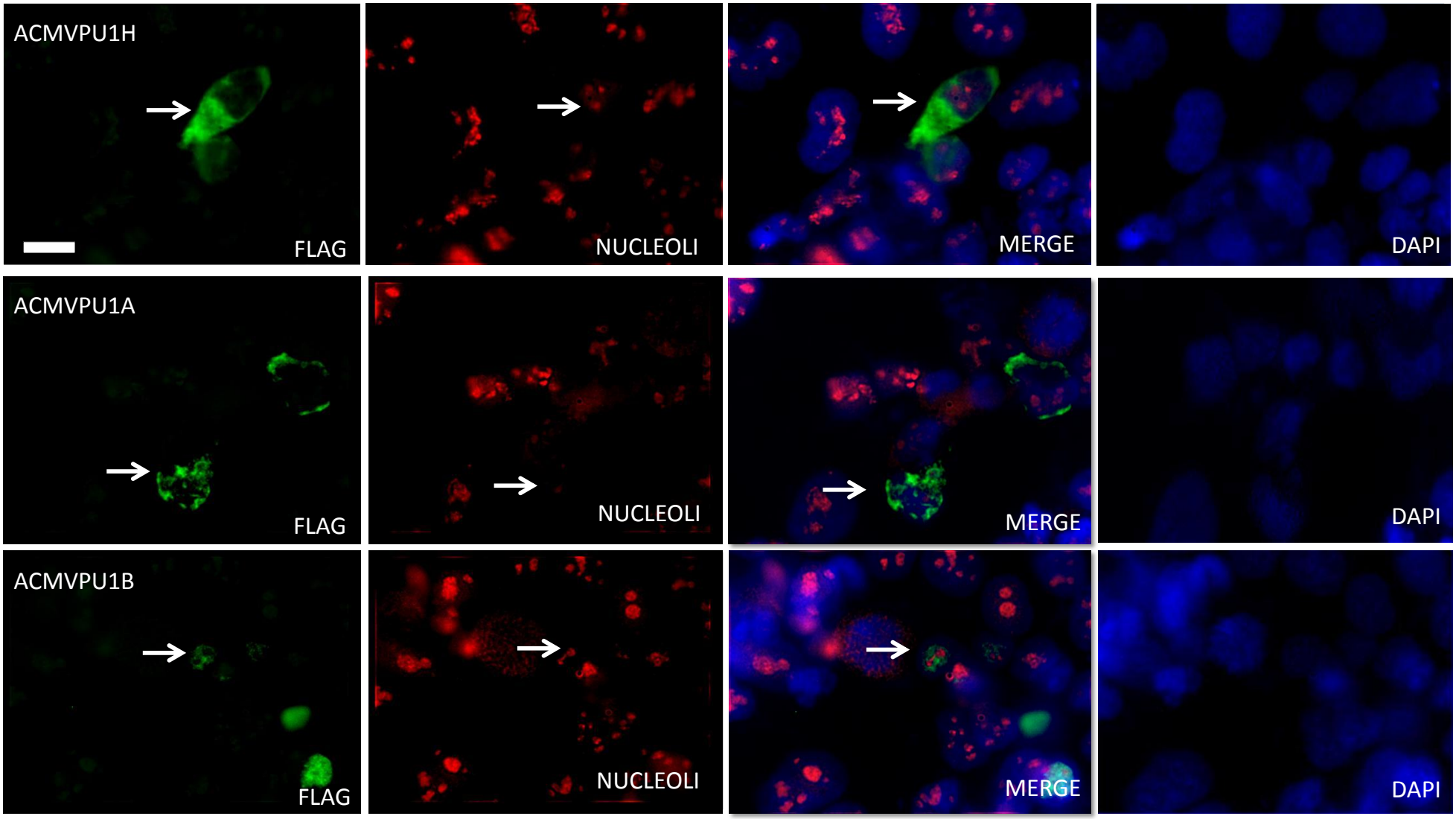
Table 4-3. Codon Analysis Table.

Four methyltransferases (M.HpaII, M.HhaI and M.SssI and the catalytic domain of DNMT3B) were codon analysed as described above. CAI values are indicators of gene adaptation with an optimum value of 1.0. For the methyltransferases analysed this value ranged between 0.63 to 0.70. Generally, the lower the CAI value, the higher the likelihood that the gene will be expressed poorly. The GC content for M.SssI was under 30% which is in the lower optimum range increasing the likelihood of a negative effect on its transcriptional and translational efficiency.

4.7 Cellular Distribution of ZFP/Methyltransferases-Immunofluorescence Analysis

Following analysis of protein expression, their cellular distribution was assessed using immunofluorescence. The method allows detection of the location and relative abundance of a protein of interest using light microscopy with a fluorescence microscope. The key to the whole procedure is the choice of an antibody capable of detecting the protein of interest (FLAG in this case) and targeting the fluorescent dye of choice to the correct antigen. I have employed secondary indirect immunofluorescence in HeLa cells with two antibodies; a primary anti FLAG® M2 monoclonal antibody (Sigma-Aldrich Catalogue No: F3165) to specifically bind FLAG-tag upon transfections with ACMVPU1 ZFP fused with the methyltransferases used in this project; M.HpaII, M.HhaI and M.SssI and a catalytic domain of DNMT3B and DNMT3A. The primary antibody was captured by a secondary anti -mouse antibody conjugated to FITC fluorophore, BD Pharmingen ™ FITC Rat Anti-Mouse IgG1 (Catalogue No: 553443) (used at 1:400 dilution), a kind gift from Dr Jonathan Morris (Cancer Cell Biology and Imaging group). FITC (Fluorescein isothiocyanate) was visualized in the green region with excitation and emission spectrum wavelengths of 495nm and 521nm respectively.

Additionally, HeLa cells were stained for Nucleolin using an anti-rabbit Nucleolin polyclonal antibody (used at 1:400 dilution) a kind gift from Professor Angus Lamond, University of Dundee. This was followed by a secondary antibody staining with Alexa Flour 633 (Life Technologies Cat No A21071) conjugated anti- rabbit antibody (used at 1:400 dilutions). The fluorophore is fluorescent in the far red region with an absorption max at 632 nm and emission max at 647 nm). The staining procedure is described in detail in the Methods Chapter 2, Section 2.13.4. Briefly, HeLa cells were transfected in a 12 well format with the ACMVPU1, ACMVPU1H, ACMVPU1/SssI, ACMVPU1/3A and ACMVPU1/3B constructs using IcaFectin™ 441 as described in Methods Chapter 2 Section 2.11.3.2. Three days following transfections, HeLa cells were trypsinized, washed once in PBS, pelleted by centrifugation at 200xg for 10 minutes and resuspended at $2.5x10^5$ cell/400µl of DMEM media supplemented with 10% FBS (v/v). Following growth on Poly-L-lysine coated slides, cells were subjected to Paraformaldehyde (PFA) fixation, Triton X-100 lysis and 10% Donkey Serum (DS) blocking (see Methods Chapter 2 Section 2.13.4.). Cells were co-stained with anti-rabbit Nucleolin polyclonal antibody (used at 1:400 dilution) and anti- FLAG® M2 monoclonal antibody (used at 1:100 dilution) overnight at 4°C. Following secondary antibody staining using fluorescent antibodies described above, the slides were DAPI stained, to allow visualization of DNA and the nucleus. DAPI is excited by ultraviolet light with absorption maximum of 358nm and emission at 461nm which appears blue/cyan. Following staining, slides were fixed with mounting solution and visualised using inverted microscopy (Zeiss microscope and AxiovertVision software version 4.3.). Results of staining are shown in Figure 4-4. A FITC FLAG-tag signal could be detected in all of the stained slides apart from the negative control, which were mock transfected cells. For the majority of the slides FITC FLAG-tag signals appear co-localized with DAPI and CY-5 labelled Nucleolin which suggests nuclear localization of the transfected constructs. Only for ACMVPU1/ HpaII did there appear to be some cytoplasmic localization in addition to nuclear. Overall it can be concluded that the NLS signals effectively drives the plasmid constructs upon transfection to its intended region which is essential for targeting DNA requirements for the ZFPs.


ACMVPU1H
FLAG
NUCLEOLI
MERGE
DAPI
ACMVPU1A
FLAG
NUCLEOLI
MERGE
DAPI
ACMVPU1B
FLAG
NUCLEOLI
MERGE
DAPI

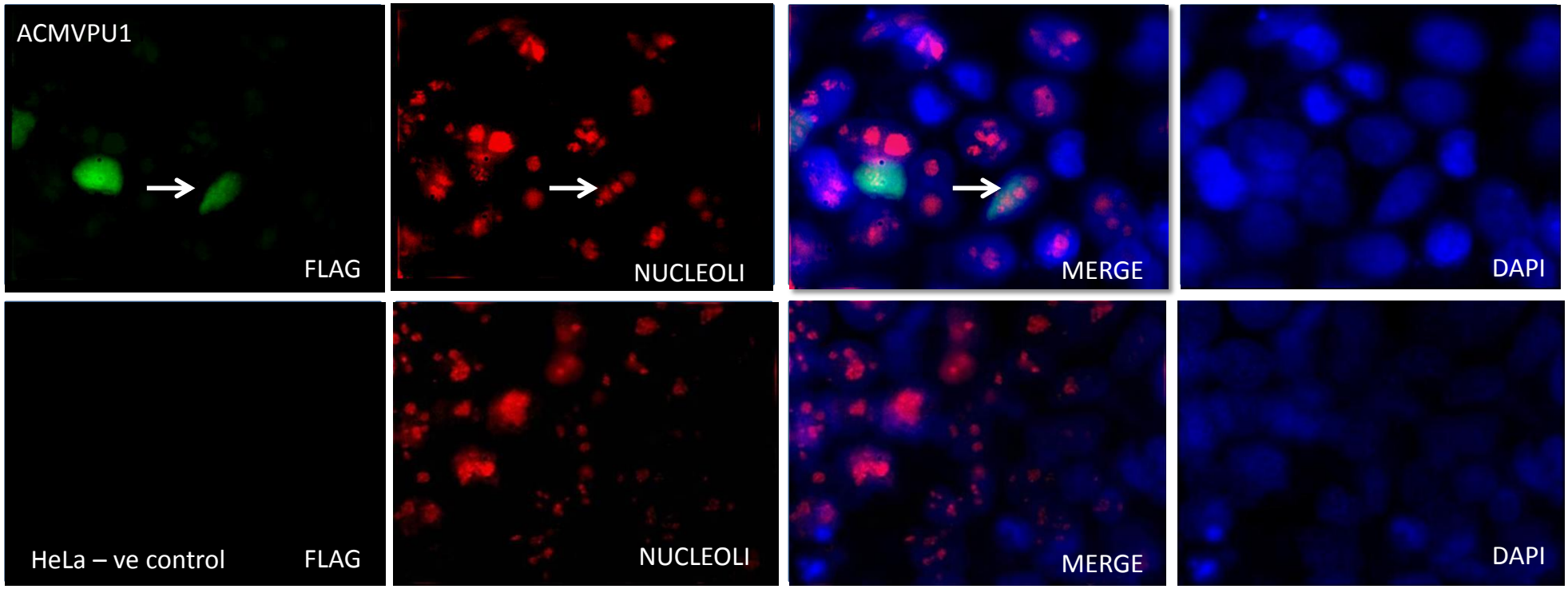


Figure 4-5 Indirect immunofluorescence analysis of FLAG-tag cellular localization.

HeLa cells were transfected with CMVPU1/Mtase/FLAG-tag or CMVPU1/FLAG-tag plasmids. Four days post transfections cells were stained for FLAG-tag using anti-FLAG mouse monoclonal antibody, followed by anti-mouse FITC conjugated secondary antibody. Cells were also stained for nucleoli using polyclonal rabbit anti- Nucleolin which was a positive control for our staining assays, followed by anti-rabbit CY5 secondary antibody. Nuclei were stained with DAPI. *Scale bar,* 10 µm.

4.8 Experimental Logistics evaluation

4.8.1 Cell Line Transfection Efficiency Evaluation

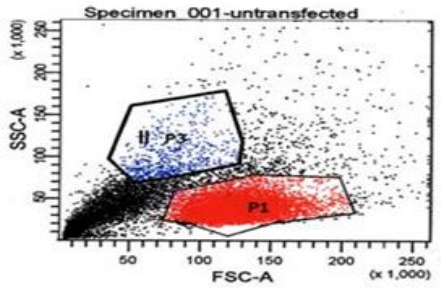
Two different cell lines were used for transfection efficiency studies. The cell line initially chosen for the purposes of the project was human promyelocytic leukaemia cell line HL60, because of its haematological lineage and because it possess a transcriptionally active *CDKN2B* gene, with an unmethylated promoter. However, as HL60 suspension cells proved difficult to transfect, the human epithelial cervical cancer HeLa cell line was also investigated as an alternative model.

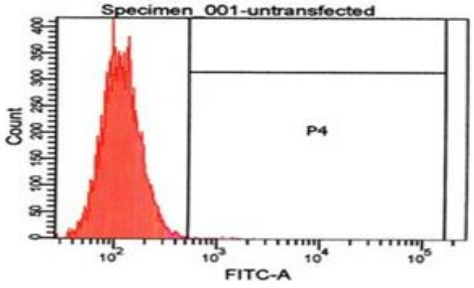
The transfection efficiency of HL60 cells was evaluated using 1µg of DNA encoding fluorescent protein GFP (pGFP Amaxa). A number of different transfection reagents were tested; a novel synthetic compound (IcaFectin™), a linear polyethylenimine derivative (JetPEI™), and nucleofection – nuclear gene transfer technology Amaxa®. The transfection protocol for each procedure is described in detail in the Methods (Chapter 2 Section 2.11.3). HL60 transfected samples were analysed 24 hours post-transfection for the presence of GFP protein using FACS (Figure 4-6).

The GFP signal was gated on mock transfected HL60 cells which were used as a negative control. The FL-1 channel was used to measure GFP intensity and FlowJo software was used to analyse transfection efficiency. Transfection efficiencies using different methods differed significantly, with nucleofection giving the best result with efficiencies of close to 70%. Nucleofection uses a device called a Nucleofector to electroporate plasmid directly into the cell nucleus with the help of cell type specific reagents and under specific electrical parameters. JetPEI™ a linear polyethylenimine derivative positively charges DNA to bind to anionic cell surface residues and enter the cell *via* endocytosis. Once inside the cell the amines are protonated resulting in an influx of counter-ions and a lowering of the osmotic potential. The resulting osmotic swelling releases the DNA:PEI complex into the cytoplasm from where the DNA diffuses into the nucleus once unpacked from the PEI complex [397]. IcaFectin™ 441 uses proprietary, synthetic molecules derived from natural compounds to form discrete compact mushroom like structures with DNA for optimal delivery into the cells.

FACSDiva Version 6.1.1

A



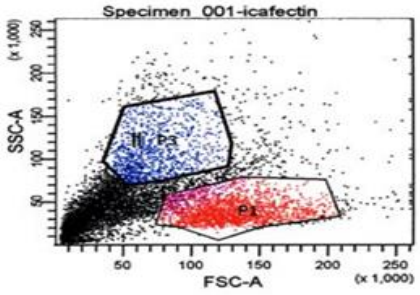


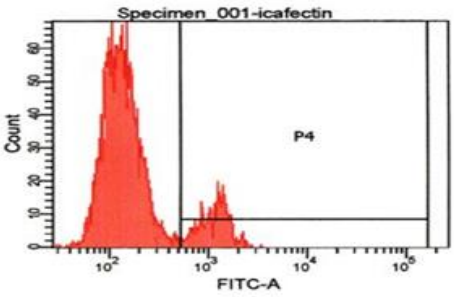
Specimen Name:	Specimen_001
Tube Name:	untransfected
Record Date:	Nov 27, 2008 12:25:52 PM
$OP:	Administrator

Population	#Events	%Parent	FITC-A Mean
P1	10,037	61.4	130
P2	57	0.6	987
P4	51	0.5	1,042
P3	629	3.8	1,318

FACSDiva Version 6.1.1

B





Specimen Name:	Specimen_001
Tube Name:	icafectin
Record Date:	Nov 27, 2008 12:35:33 PM
$OP:	Administrator

Population	#Events	%Parent	FITC-A Mean
P1	2,071	22.0	302
P2	311	15.0	1,205
P4	310	15.0	1,207
P3	1,043	11.1	1,206

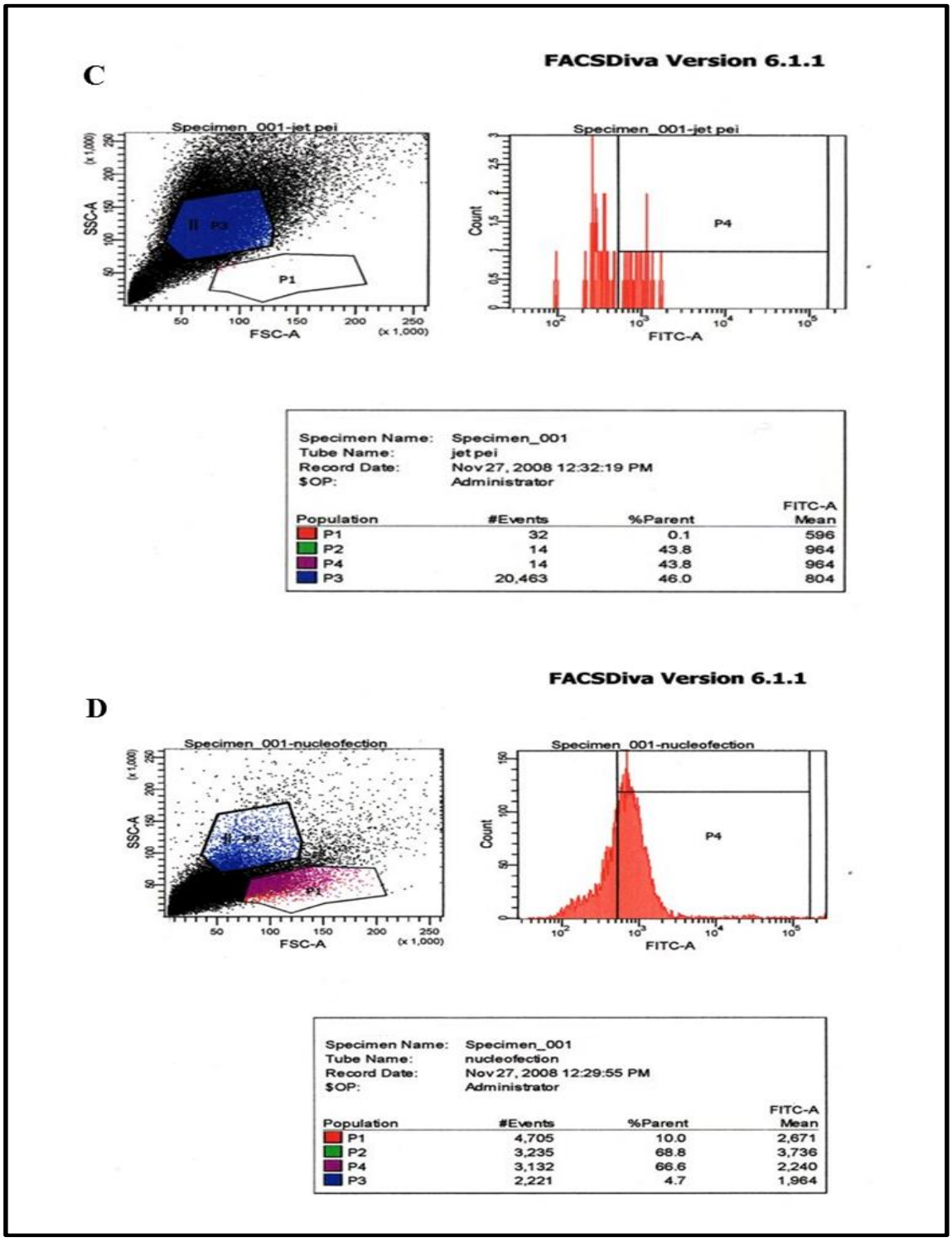


Figure 4-6 HL60 Transfection efficiency.

pGFP Amaxa transfected HL60 cells were assayed for GFP expression 24 hours post transfection using three different methods; Amaxa® Nucleofection, IcaFectin™ and JetPEI™ **A** Untransfected HL60 cells. **B** IcaFectin™ transfection efficiency was 15%, and the cell population appears healthier compared to nucleofected samples (**D**). **C** JetPEI™ reagent appeared to have toxic effect on HL60 cells. **D** The best transfection efficiency was achieved using Amaxa® nucleofection (66%).

In subsequent nucleofection experiments, low HL60 cell viability was observed. The cells appeared morphologically different compared to un-transfected HL60 cells, upon light microscope observation and failed to proliferate further. This was assessed by Trypan blue exclusion and cell counting at daily intervals for four days following nucleofection (see Figure 4-7).

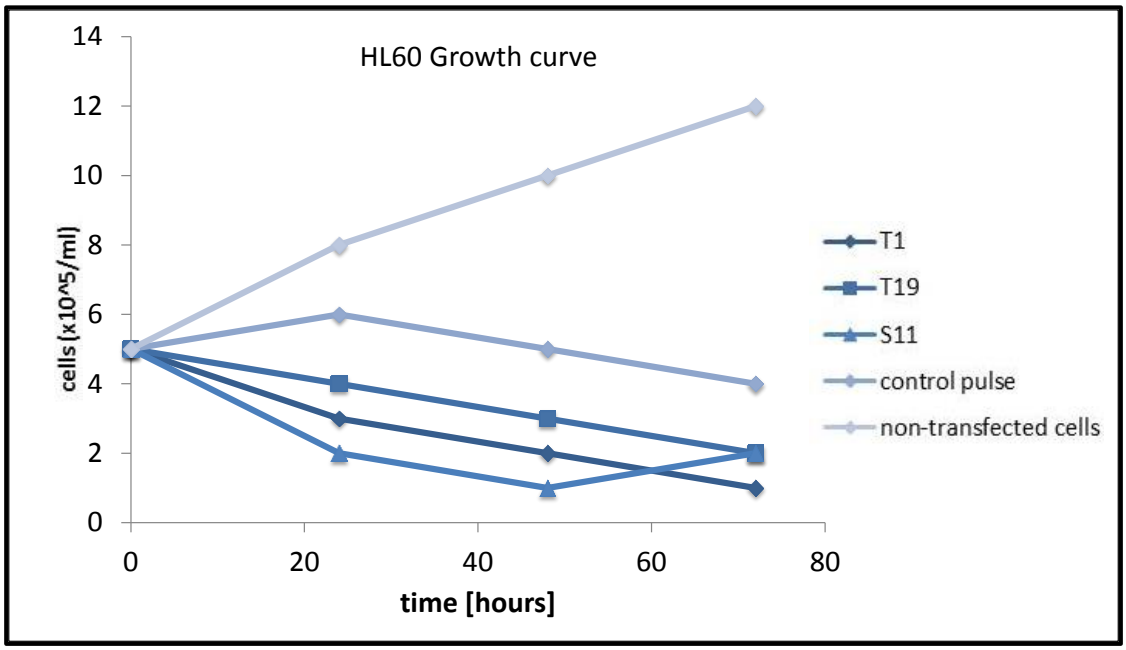


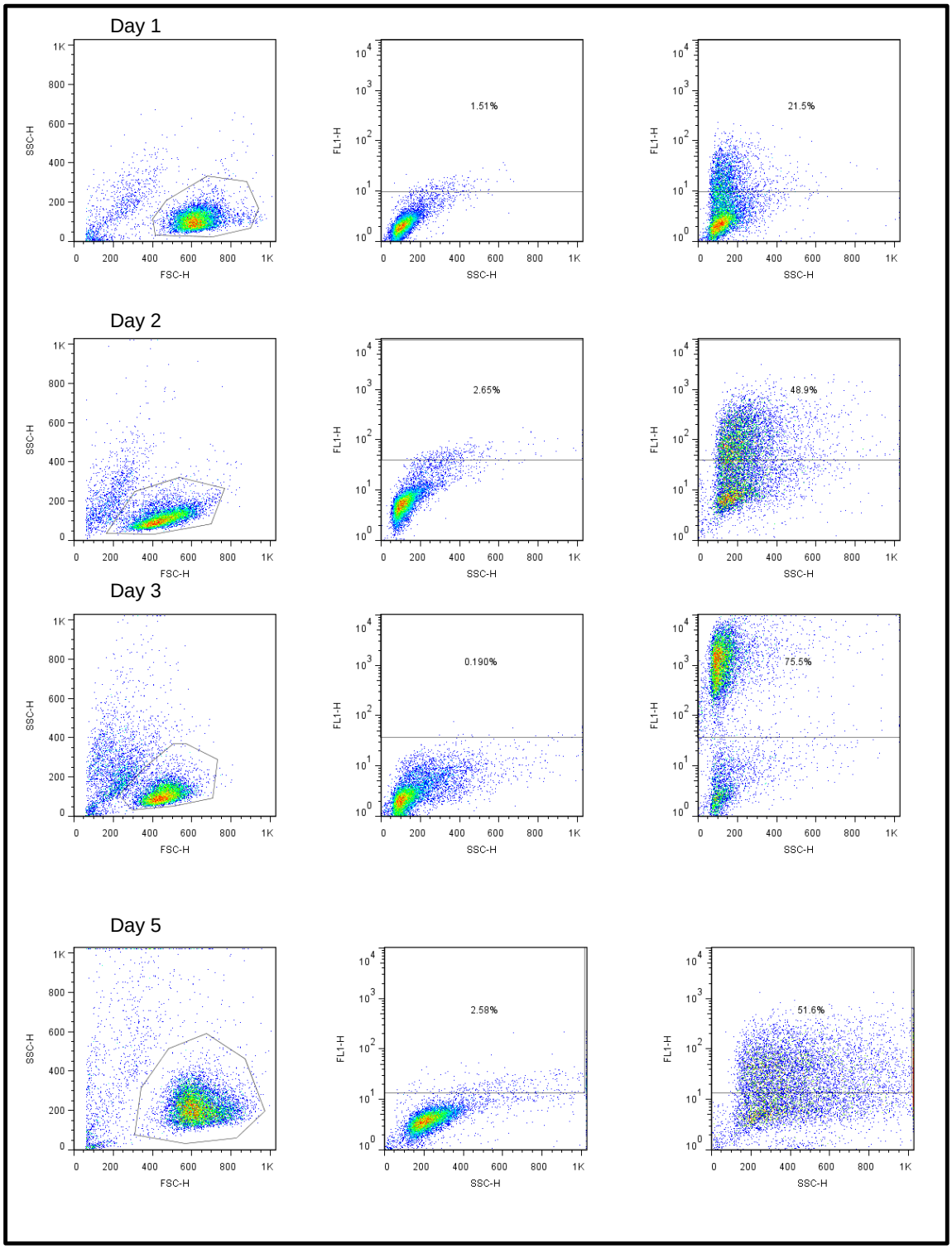
Figure 4-7 HL60 Nucleofection efficiency Assessment

HL60 cells were nucleofected as previously described using a pGFP Amaxa plasmid. Cells were nucleofected using different Amaxa programmes either for high transfection efficiency (T1 and T19) or high post-transfection cell viability (S11). GFP transfection efficiency was assessed 24h post-transfection using a fluorescent microscope with 60% cells positive for GFP expression. Cells were counted using Trypan blue exclusion at 24 hour intervals for four days. Irrespective of the nucleofection programme used, HL60 cells failed to proliferate further. Similar experiment was conducted previously and HL60 nucleofection programmes were selected on the basis of that paper [398].

As an alternative strategy to HL60 transfection lentiviral transduction approach on HL60 cells was investigated. Standard lentiviruses are capable of stable integration of their genome into the host genome. Because the goal was to deliver a transient promoter targeted DNA methylation event, I chose to explore the possibility of using second generation, non-integrating lentiviral vectors (NILVs) which harbour integrase mutations that reduce the likelihood of stable transgene integration.

A series of plasmids were used to produce the lentiviruses in Human Embryonic Kidney 293 cells (293T) cell line with a biotinylated VSV-G envelope and included a lentiviral expression plasmid containing the GFP gene inserted between the lentiviral LTRs, a packaging plasmid, (pLV-HELP), encoding the *pol, gag, rev* and *tat* viral genes and containing the *rev*-response element (RRE); and a pseudotyping plasmid pLV-iVSV-G, encoding the G protein of the Vesicular Stomatitis Virus (VSV-G) envelope gene. The plasmids were generously provided by Dr Lucas Chan of the Department of Haematological and Molecular Medicine. Lentiviral particles were concentrated using streptavidin conjugated magnetic beads according to methods published by Dr David Darling of this department [399] and under the kind guidance of Dr Darling the particles were used to infect HL60 cells. Upon target cell infection, the viral RNA is reverse-transcribed, and imported into the nucleus.

GFP expression was assessed at three day intervals for up to 18 days by FACS and analysed using FlowJo software Figure 4-8. Although significant transduction efficiency was obtained using the lentiviral transduction method, even after 18 days too high a percentage of cells remained GFP positive, suggesting stable integration of the GFP gene. This was deemed unacceptable as we required establishment only of transient expression of targeted methyltransferase enzymes.


Day 1
SSC-H
FSC-H
FL1-H
SSC-H
1.51%
21.5%
Day 2
2.65%
48.9%
Day 3
0.190%
75.5%
Day 5
2.58%
51.6%

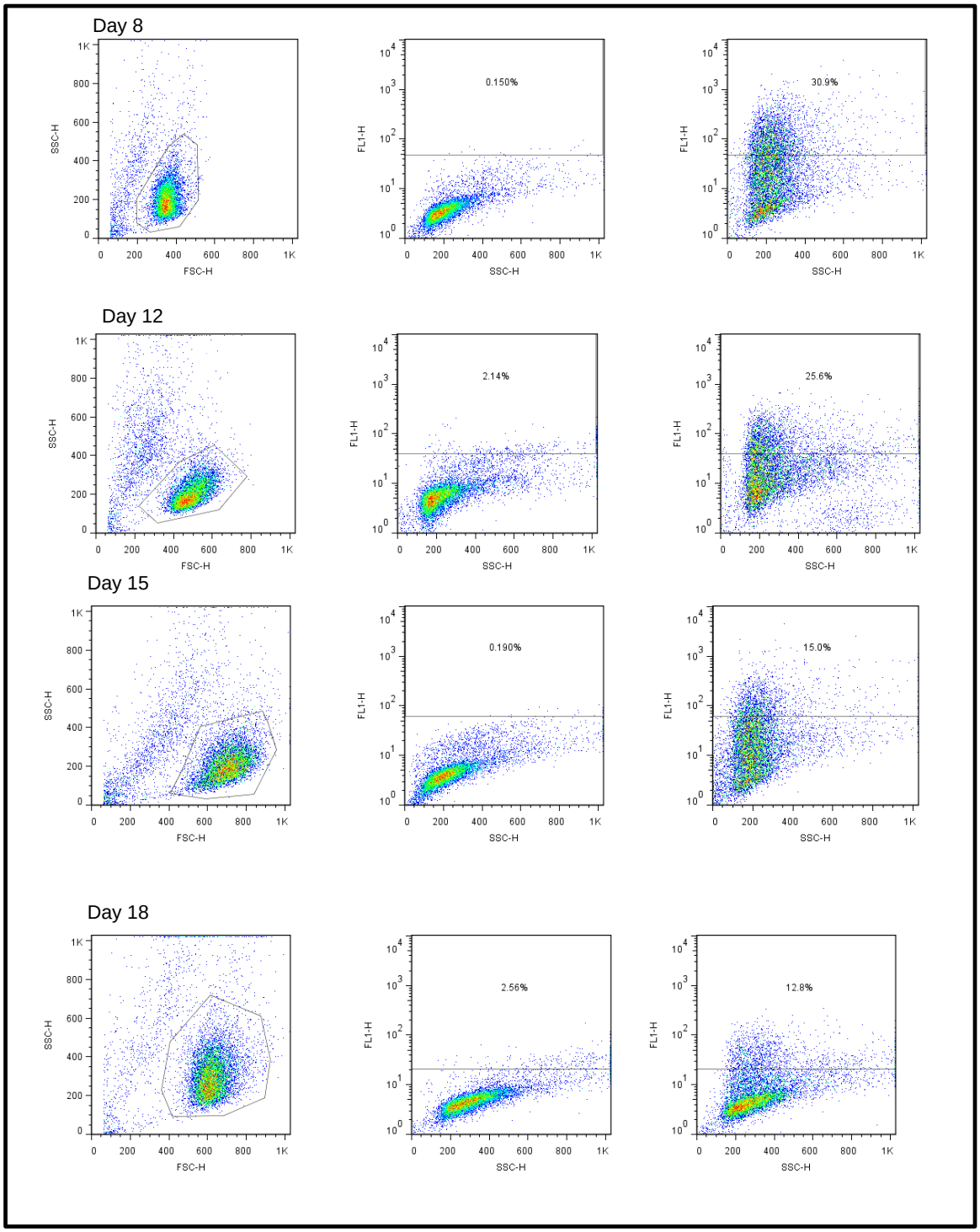


Figure 4-8 GFP expression in HL60 following Lentivirus transduction.

Expression was monitored for 18 days following virus transduction by FACS. Best transfection efficiency was achieved at Day 3 of transfection (75.5%). The GFP signal was detected as far as Day 18.

Given the inability to either efficiently transiently transduce HL60 cells using lentiviral particles or generate proliferating cells following nucleofection, the HeLa cell line was subsequently evaluated as an alternative system. This is an adherent cell line easily transfectable, receptive to plasmid vectors and widely used for such purposes. Additionally, the *CDKN2B* promoter is unmethylated in this cell line, and $p15^{INK4B}$ protein is expressed.

Transfection efficiency and cell viability following HeLa transfections, using Icafectin 441™, CaBES and Amaxa® nucleofection were assessed as described above for HL60 evaluations. HeLa pGFP Amaxa transfected samples were analysed 24 hour post-transfection for the presence of GFP protein using FACS as previously described. Transfection efficiencies using different methods were consistently very good, at around 70% for all the transfection methods, as shown in Figure 4-9.

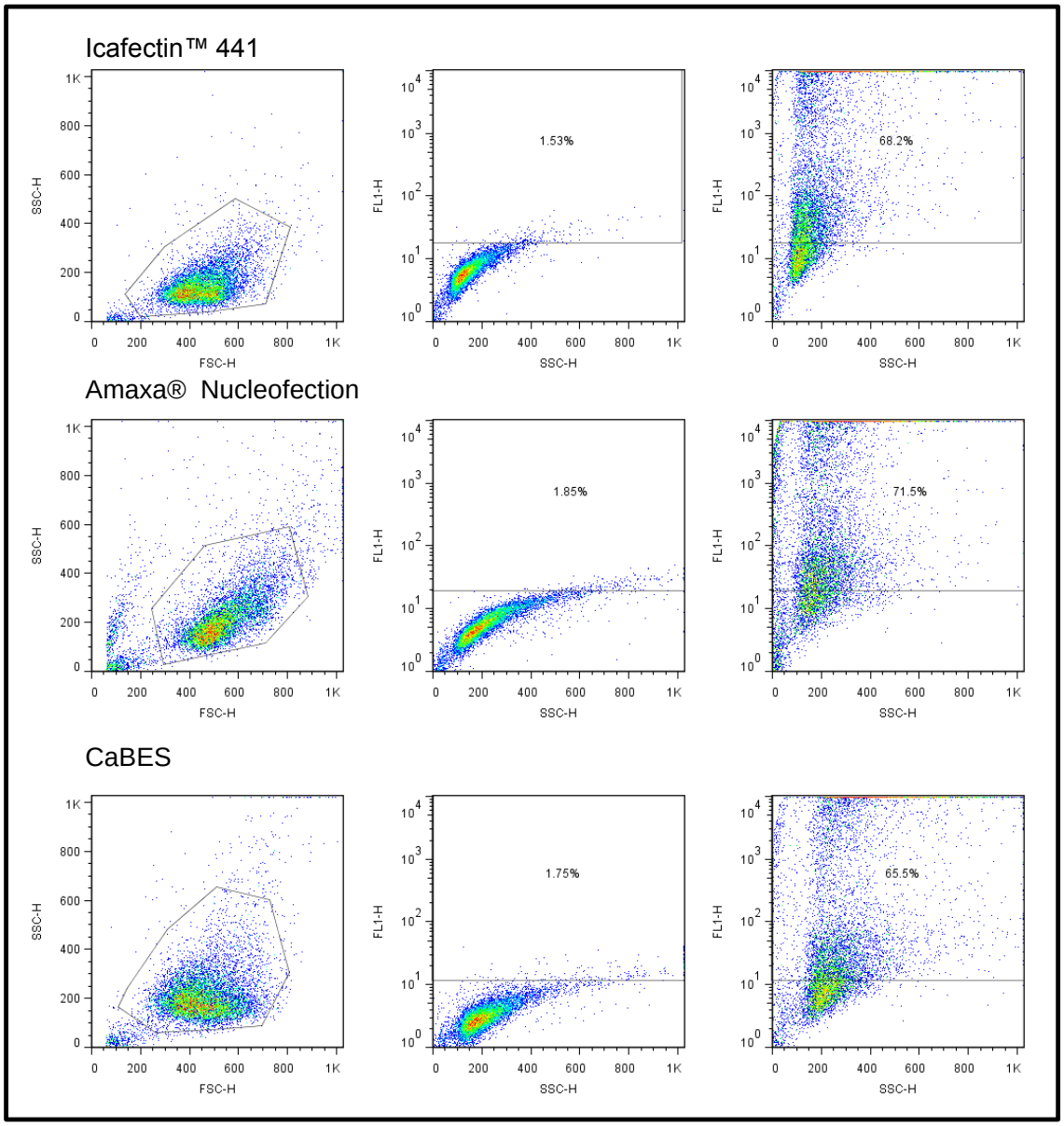


Figure 4-9 HeLa Transfection efficiency evaluation

Measurement of GFP expression as an indicator of transfection efficiency of HeLa cells. HeLa cells were transfected with pGFP Amaxa using three different methods; Amaxa® nucleofection, Icafectin™ 441 DNA transfection reagent and CaBES method. Best transfection efficiency (71.5%) was achieved using Amaxa® nucleofection method, followed by Icafectin 441™ (68.5%) and CaBES (65.5%).

Cells were also monitored for their proliferation rate, recovery and viability by Trypan Blue exclusion as shown in Figure 4-10. Nucleofected cells appeared to take longer to recover (3 days), compared to Icafectin™ 441 and CaBES transfections. CaBES was initially selected as the optimal transfection method due to high transfection efficiency, good cell viability, ease of use and costs.

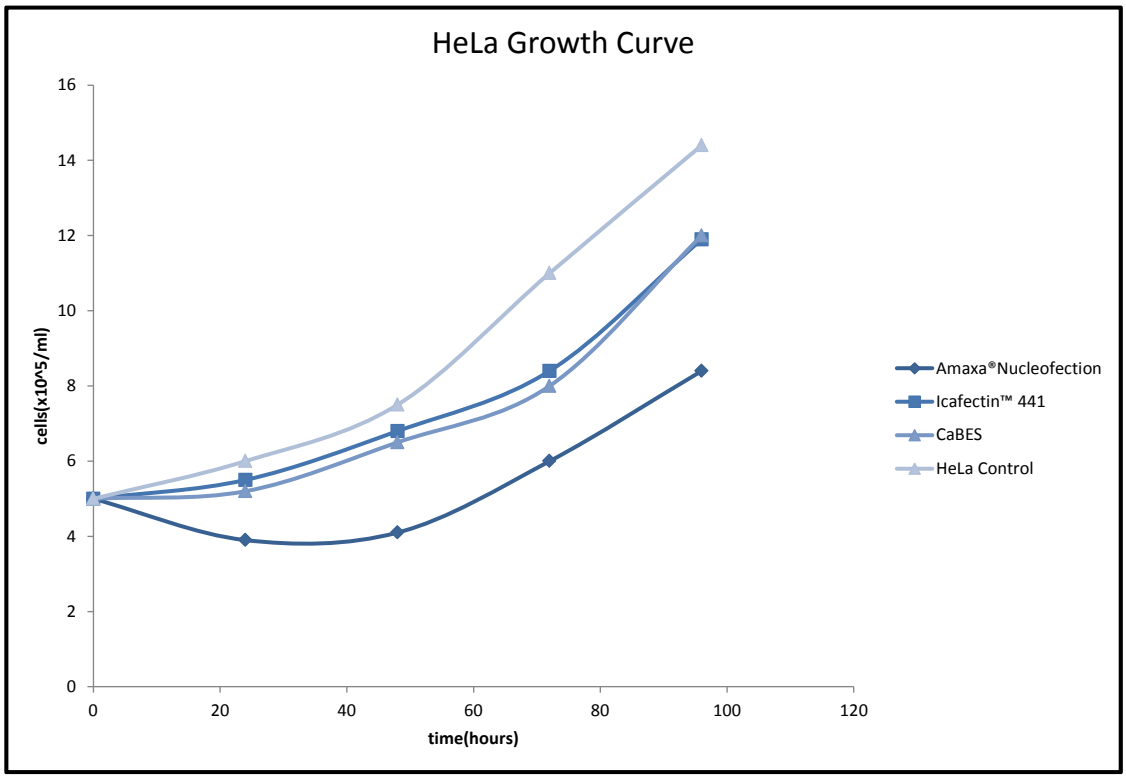


Figure 4-10 HeLa cell proliferation assay

HeLa cell viability was assessed using Trypan Blue exclusion, following Amaxa® nucleofection, Icafectin™441 or CaBES transfection. These transfection procedures are described in detail in Methods Chapter 2, Section 2.11.3. Following transfections, HeLa cells were counted at 24 hour intervals for up to 96 hours post transfections and were found to be proliferating and viable.

4.9 *In vivo* targeting of methylation to the *CDKN2B* promoter

Having established optimal transfection conditions and generated constructs expressing ZFP/ methyltransferase fusion proteins, I proceeded to an *in vivo* assessment of *CDKN2B*-targeted methylation. Targeting DNA methylation to the *CDKN2B* promoter region of HeLa cells involved co-transfection of cells with 10µg of plasmid DNA (various ZFP/methyltransferase constructs) spiked with 0.5µg of Amaxa™ GFP vector, using the CaBES method described previously.

To determine expression of FLAG-TAG - ZFP/methyltransferase fusion proteins (outlined in Table 4.2.) samples of transfected HeLa cells ($1X10^6$) four days post transfection were western blotted and probed using anti-FLAG Clone M2 monoclonal antibody (Sigma-Aldrich Catalogue Number F3165) as described in Section 4.6. above and shown in Figure 4-2.

Transfected HeLa cells were routinely passaged and cell samples taken for the following analysis: Western Blot ($1X10^6$) for analysis of $p15^{INK4B}$ protein expression, *CDKN2B* gene transcriptional status by Taqman™ RT-qPCR ($1X10^6$), and DNA methylation analysis using Combined Bisulfite Restriction Analysis COBRA , and/or bisulfite sequencing. Therefore, samples taken at various time intervals, namely Day 4, 7, 10, 14 and 21, were used to monitor *CDKN2B* promoter methylation and transcriptional/translational status of the gene.

4.9.1 Results of *in vivo* targeting of *CDKN2B*

One of the primary aims of this project was to assess if it possible to target methylation to the endogenous promoter of an actively transcribing gene. As this had never been attempted before to our knowledge, the exact effects of such targeting were unknown, but it was hoped that the transcriptional regulation of the targeted gene would perhaps be affected due to the interference with transcription factor binding, access of transcriptional machinery or methyl-CpG binding domain protein (MBD) mediated recruitment of histone deacetylases and other chromatin re-modellers, leading to the formation of an inactive heterochromatin state as discussed in detail in the Introduction (Chapter 1) of this thesis. This assumption is based on the previous observations of correlation between methylated promoters and respective gene transcriptional status [400]. Once laid down, DNA methylation could potentially lead to recruitment of other epigenetic co-factors such as maintenance methyltransferases and histone deactelyase, to provide a framework for a self- reinforcing propagation of a methylated state [401]. It was previously described that the DNA methylation density may also influence the stability of the epigenetic imprint.

A second important issue to be addressed by this project was the effect of low density vs. high density methylation and the spread of methylation within the *CDKN2B* promoter. Studies on murine erythroleukemia cells containing a chromosomally integrated proviral reporter, premethylated at different densities, have shown a faithful propagation of high density of methylation compared to low level of methylation which was unstable, and accompanied by eventual complete demethylation [402]. To address this issue I have used prokaryotic (M.HpaII, M.HhaI) methyltransferases which have only a few target sites within the *CDKN2B* promoter, to lay down low density methylation and M.SssI, and the catalytic domain of eukaryotic methyltransferases DNMT3A and DNMT3B to introduce high density methylation levels (see Figure 4.11)

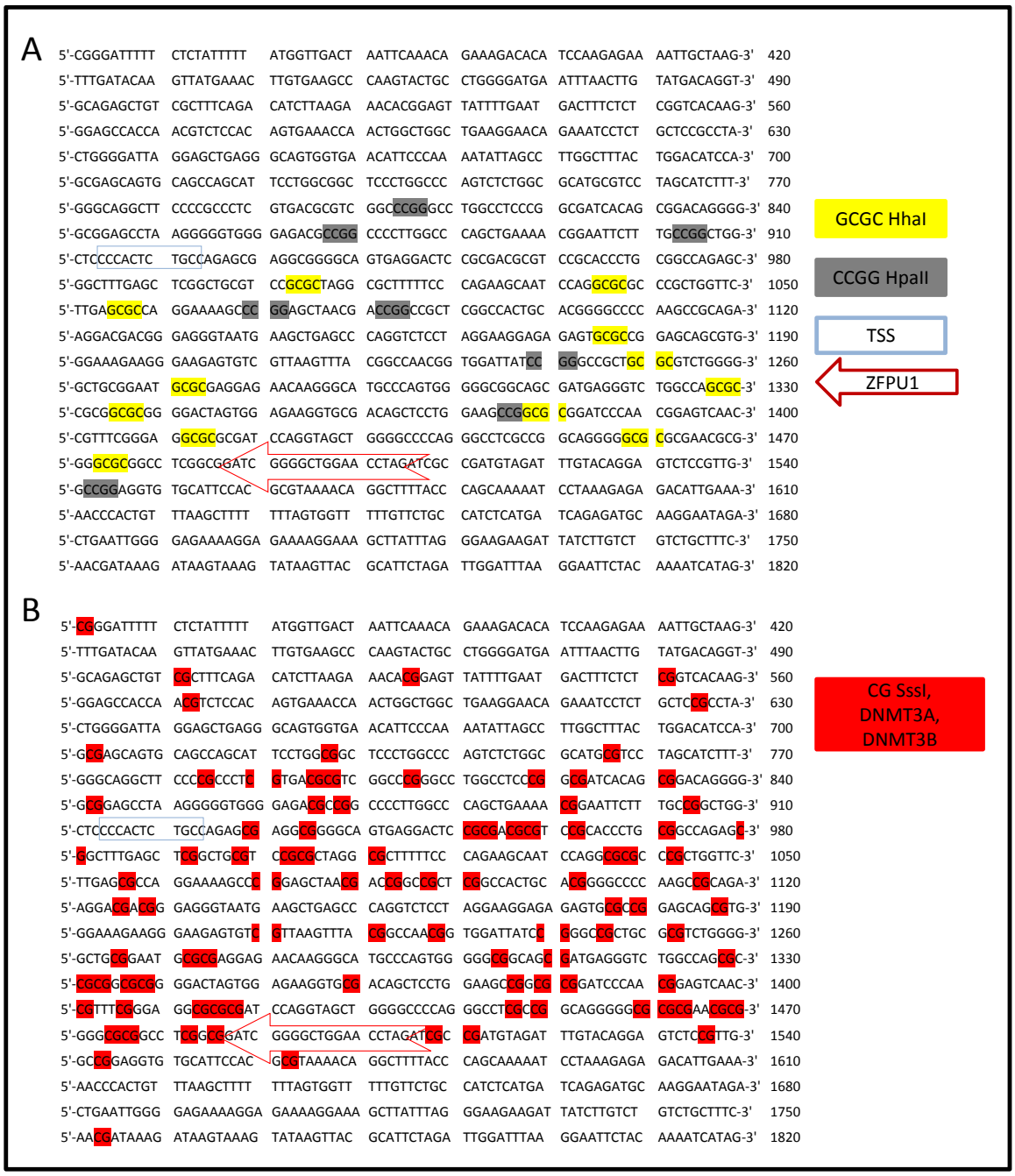


Figure 4-11 Methylation density possibilities at the *CDKN2B* promoter for various targeted methyltransferases.

Representative diagram of promoter methyltransferase target sites within the *CDKN2B* promoter. **A** There are seven HhaI sites –target sequence GCGC and four HpaII sites –target sequence CCGG upstream of the most 3' zinc finger target binding site. Targeting *CDKN2B* using ZFPs fused with M.HpaII or M.HhaI allows for relatively low density methylation seeding in comparison to **B.** There are potentially 64 CG sites available for methylation using ZFPs fused with M.SssI or the catalytic domains of DNMT3A, DNMT3B methyltransferases.

Assessment of successful targeted methylation of the *CDKN2B* promoter, and its effect on the levels of expression/translation of the $p15^{INK4b}$ involved analysis of samples of HeLa cells at set time points; Days 4,7,10, 14 and 21 post-transfection.

The first step was to evaluate the efficiency of targeted methylation at the *CDKN2B* promoter using ZFP/methyltransferase fusions. Samples were taken (at the above specified time points) from $5x10^5$ HeLa cells, genomic DNA prepared and bisulfite converted. Bisulfite conversion involves deamination of unmethylated cytosine residues to uracil, but leaves 5-methylcytosine unaffected (Figure 4.12A). In this way specific methylation patterns within a DNA sample can be investigated.

The region of interest in the *CDKN2B* promoter was PCR amplified using primers not biased towards methylation, by semi-nested PCR (- 43 to +268 bp relative to TSS). PCR samples were subjected to methylation status analysis, using Combined Bisulphite Restriction Analysis (COBRA), followed by bisulfite sequencing. COBRA is a semi- quantitative method that utilizes restriction enzymes to selectively distinguish between methylated and non-methylated CpGs [403].

To test the robustness of the COBRA assay, HL60 and Raji cell lines were examined. The HL60 cell line has previously been shown to have an un-methylated *CDKN2B* promoter [404, 405], and express $p15^{INK4B}$ [406]. This cell line is established from human promyelocytic leukaemia cells and is known to spontaneously differentiate into a number of different cell types such as neutrophils or mature granulocytes with distinct gene expression profiles during routine culturing [407] [408]. To ascertain if differentiation might affect the methylation status of the *CDKN2B* gene promoter in HL60 cells during differentiation, cells were cultured in the presence of the differentiating agent DMSO (1.5%). Differentiation was confirmed by flow cytometric measurements of CD11b, a well-known early differentiation marker. The Raji cell line is typically used as a positive control for methylation analysis, as the *CDKN2B* promoter is known to be heavily methylated in this cell line [409]. Genomic DNA was extracted from the above cell lines, bisulphite converted and the region of interest (- 215 to +410 bp relative to TSS) amplified using PCR. This was followed by a second round of semi-nested PCR to increase the specificity of the PCR product.

The product was then digested using *R.BstUI* restriction enzyme whose recognition site is left intact if the DNA is methylated at both cytosines, but converted to TGTG if unmethylated. In order to check for bisulphite conversion efficiency another enzyme, *R. StyI*, which recognizes the sequence CCTTGG, was used. Successful bisulphite conversion should completely destroy this site and the DNA will not be cleaved. The specifics of bisulfite conversion and the COBRA assay are outlined in Figure 4 -12. below.

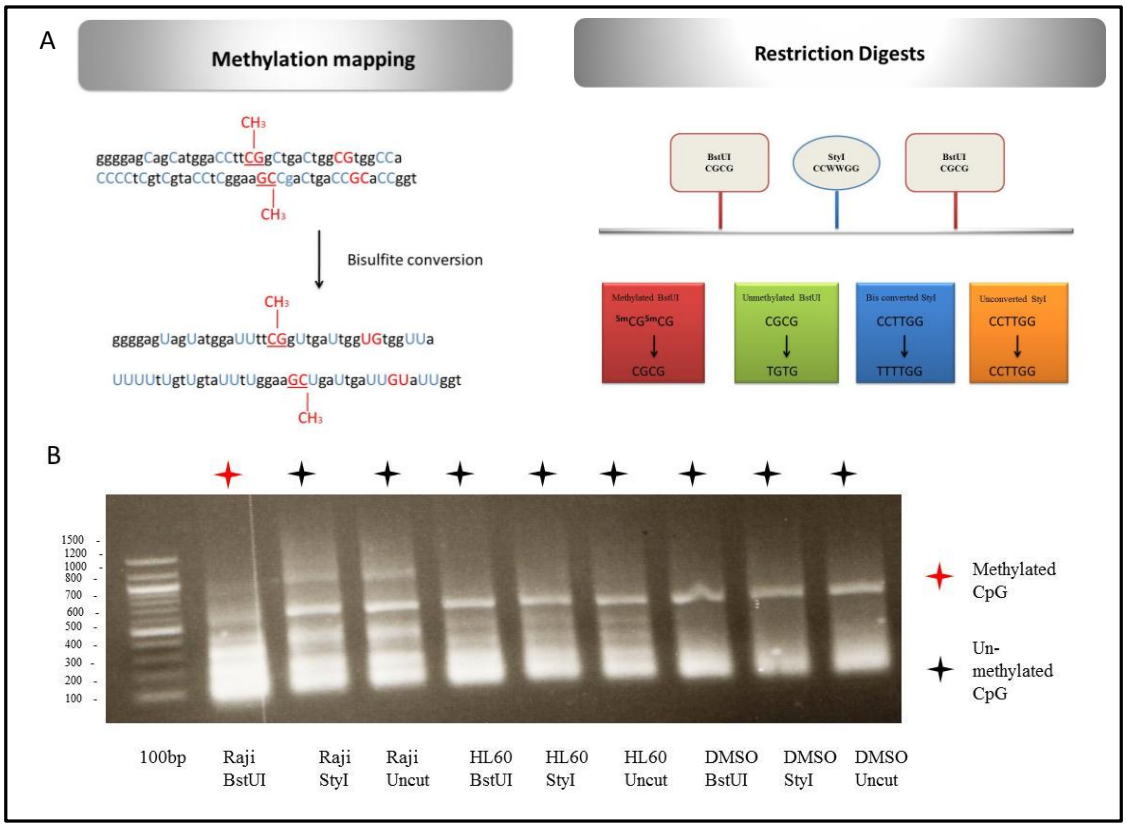


Figure 4-12 COBRA Assay

Schematic representation of the assay principle showing **A** bisulfite conversion of cytidine (C) to uracil (U) unless the cytidine is methylated within CG context, in which case the conversion will not occur. **B** Bisulfite converted DNA was amplified in semi-nested PCR to generate a PCR product 601bp in size. This product was then subjected to restriction analysis using *R.BstUI* (CGCG) which cleaved only the methylated product Raji (arrowed), as expected whilst HL60 with or without DMSO remained uncut. It can be concluded that the *CDKN2B* region scrutinized in HL60 is unmethylated irrespective of cell lineage whilst Raji has methylated *CDKN2B*. The conversion control was a digest with *R.StyI,* which remained uncut indicating that the CCTTGG sequence within the promoter was successfully converted to TTTTGG which could then not be cleaved with this enzyme. Enzyme efficiency was confirmed using control plasmid DNA (data not shown).

4.9.1.1. Low density *de novo* methylation at the *CDKN2B* promoter is maintained through successive cell divisions

For this aspect of the project ZFs fused with M.HpaII and M.HhaI prokaryotic methyltransferases were used to target low density methylation to the *CDKN2B* promoter region. HeLa cells were transfected with ACMVPU1H, ACMVPU1Hh and ACMVPU1 (negative control). Transfections were carried out as previously described and the samples were collected for methylation status analysis as described in the Section 4.9.1. at days 4,7,14 and 21. Genomic DNA was prepared and bisulfite converted as described previously.

Primers for COBRA analysis and bisulfite sequencing were designed spanning a region -212 bp to + 260 bp relative to the TSS to capture as many M.HpaII (CCGG) and M.HhaI (CGCG) sites as possible. Once generated, PCR products were restriction digested using *R.DpnII* for ACMVPU1H targeted DNA, and *R.BsmI* for ACMVPU1Hh targeted DNA. The *R.DpnII* recognition sequence is GATC. Only methylated C would remain following bisulfite conversion allowing for the PCR product to be cleaved, whilst unmethylated C in this sequence would be converted to T and the PCR product could not be cut by this enzyme. Similarly, *R.BsmI* has a recognition sequence GAATGCN and only methylated C can remain intact following bisulfite conversion, so the consequences for this restriction site will be the same.

To show that the methylation was due to specific targeting of HpaII deposited at *CDKN2B* rather than a non- specific effect due to presence of methyltransferase an additional control, ArbFH was used. This ZF recognition site was not present in the *CDKN2B* gene promoter sequence. This control was used to demonstrate that the methylation observed was not just a random event that can be attributed to the methyltransferase activity methylating indiscriminately.

Figure 4-13 shows the result of restriction analysis using the enzymes discussed above on the samples taken at various time points. To confirm bisulfite conversion had taken place *R.Tsp5091* was used as a conversion control check. The recognition sequence for this enzyme is AATT and is present at positions 34 and 168 of the amplicon only in the bisulfite converted DNA. Therefore, bisulfite converted DNA would yield two bands 134bp and 338bp in size only if successfully bisulfite converted. *R. DpnII* enzyme cuts at the underlined sequence within *CDKN2B* only if the highlighted C within TCGG (M.HpaII target site) context is methylated AACGGAT**C**GGTCG. *R.DpnII* restriction yields two distinct bands of 219bp and 253bp in size only in methylated products (asterisked in red). Similarly for ACMVPU1Hh targeted sample a restriction with *R.BsmI* recognizes and cleaves only methylated (highlighted in red) region GCGGAAT**GCGC**AGG to give two distinct bands 413bp and 59bp in size (asterisked in black).

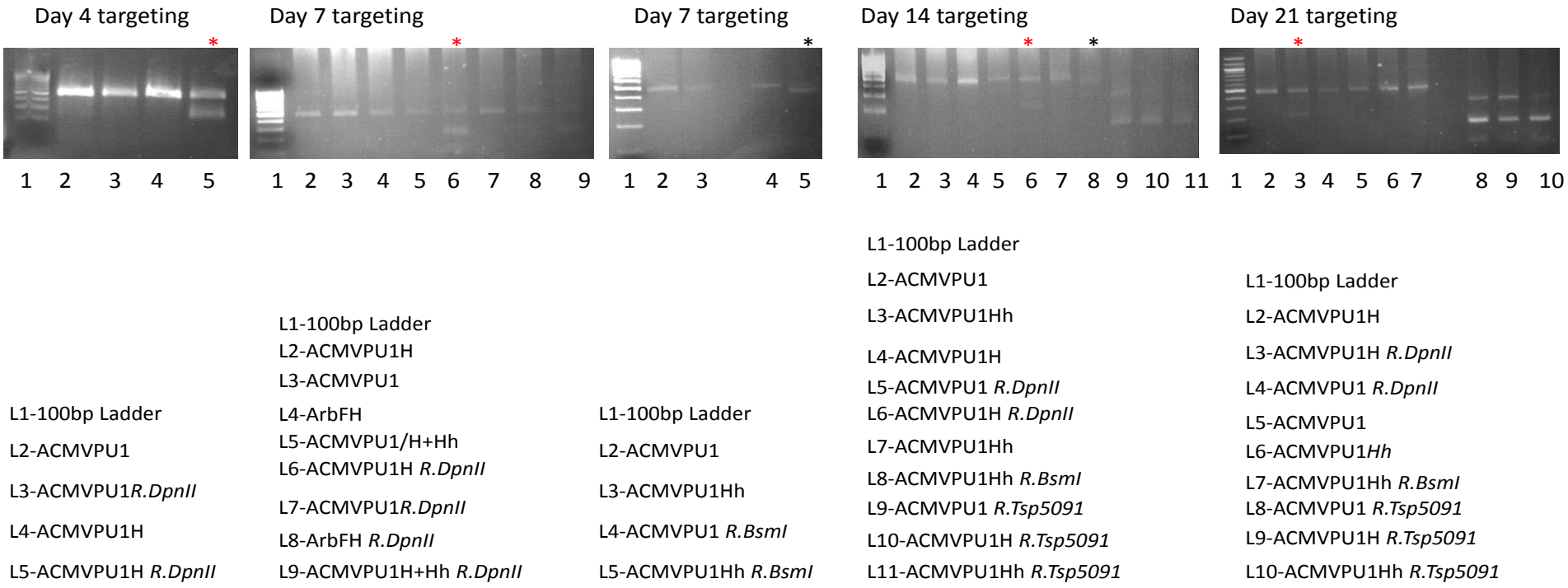


Figure 4-13 Methylation Analysis - COBRA Assay

Restriction digests of PCR products obtained from samples following targeting with ACMVPU1H, ACMVPU1Hh and ACMVPU1 constructs. To ascertain whether methylation occurred using ACMVPU1H fusion, prepared PCR products were subjected to *R.DpnII* restriction. Negative control used throughout all of the analysis procedures was a ZFP on its own, ACMVPU1, which due to the fact that it has no methyltransferase component fused to it, is expect to bind to the target region but not to methylate it.

As can be seen in the above figure, targeted methylation *via* expression of ACMVPU1H and ACMVPU1Hh was successful. The *de novo* methylation was observed four days following transfection, and persists through to Day 21 for the targeted HpaII methyltransferase. For M.HhaI constructs, methylation was observed at Day 7 and 14 but not Day 21. To confirm the findings I have selected Day 4 and Day 21 sample time points for bisulfite sequencing.

Methylation within the *CDKN2B* promoter at Day 4 for ACMVPU1H was confirmed relative to control DNA and is shown in Figure 4-14 below. The figure shows sequencing data for ACMVPU1H vs. ACMVPU1 and the methylated TCGG vs. unmethylated TTGG HpaII target sites. The M.HpaII target sequence is CCGG and only the internal cytosine residue becomes methylated [410]. This gives TCGG following bisulfite conversion. If the site is unmethylated both cytosines will be converted to thymidine. Methylation was observed at HpaII sites 300bp to 579 bp away from the ZF target site. There were four sites nearer to the ZF site and they remained unmethylated (see Figure 4-17). The implication of this will be discussed later on in this Chapter.

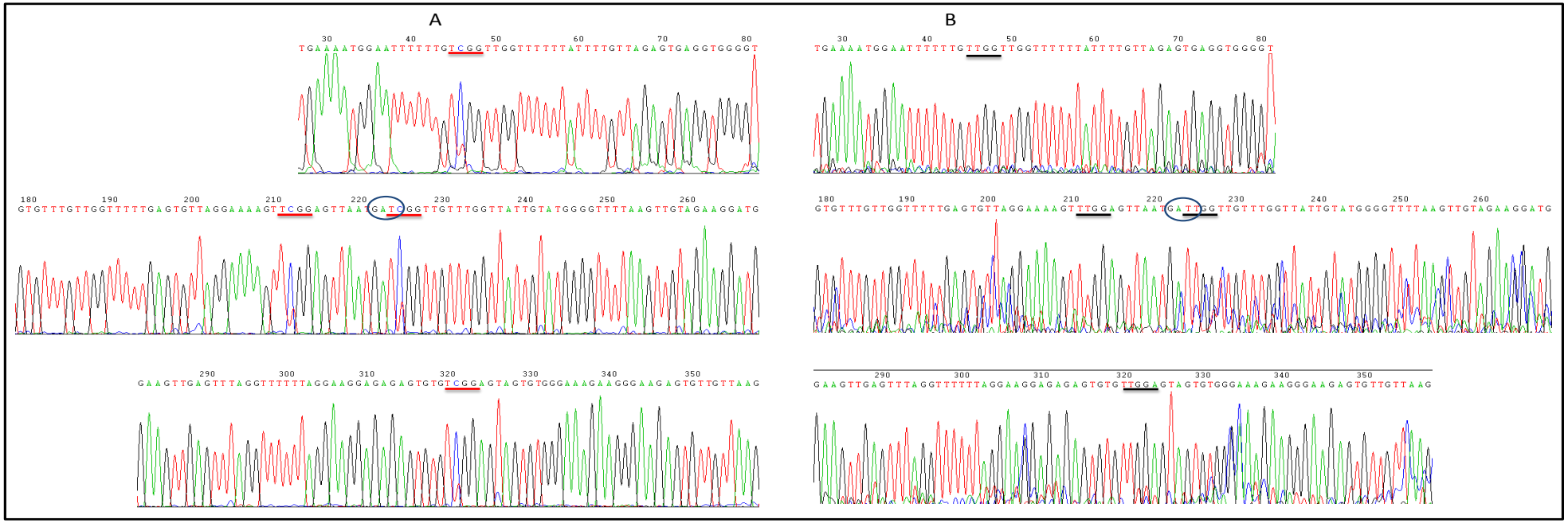


Figure 4-14 Methylation Analysis – Bisulfite Sequencing

PCR products obtained from library samples at Day 4 following targeting with ACMVPU1H, and ACMVPU1 constructs were gel cleaned as described in Chapter 2, Methods Section 2.8.2. To confirm methylation observed by COBRA (shown in Figure 4-13) ACMVPU1H and ACMVPU1 transfected samples, prepared in the same way as for COBRA analysis were bisulfite sequenced commercially. **A** The underlined sequence within *CDKN2B* highlighted in red shows methylated C within TCGG (M.HpaII target site) context. There were eight M. HpaII target sites (CCGG) available in the region upstream from the PU1 ZFP binding site. This PCR amplicon covers four of them and they were all found to be methylated. Highlighted in blue are all the potential methylatable CpG sites available for endogenous *de novo* methylation. **B** Same region is shown for the ACMVPU1 transfected cells with unmethylated target site shown underlined in black. This region shows no evidence of methylation. *R.DpnII* recognition sequence (GATC) used for restriction analysis is circled.

The above samples were actually a library of different sequence clones. Individual sequences were TOPO cloned and analysed individually. The bisulfite converted samples at time-points Day 7 and Day 21 were PCR amplified by a semi-nested approach as described previously and subcloned into the TOPO® cloning vector. DNA was subsequently extracted from eight individual clones and sequenced for methylation pattern assessment. As shown in Figure 4-15, promoter methylation is observed at Day 7 and actively maintained through following successive cell divisions to Day 21. The results indicate that there is no spread of methylation to surrounding CpGs within this region as only HpaII sites were seen to be methylated at Day7 or 21 (data not shown, see Appendix Supplementary Data CD). For the region analysed the first methylated HpaII site was observed upstream at -7 from the TSS. Three HpaII sites located downstream at positions +148, +157 and +249 relative to TSS were also methylated. However, even at Day 21, no methylation at any of the other potentially available 24 CpG sites was detected.

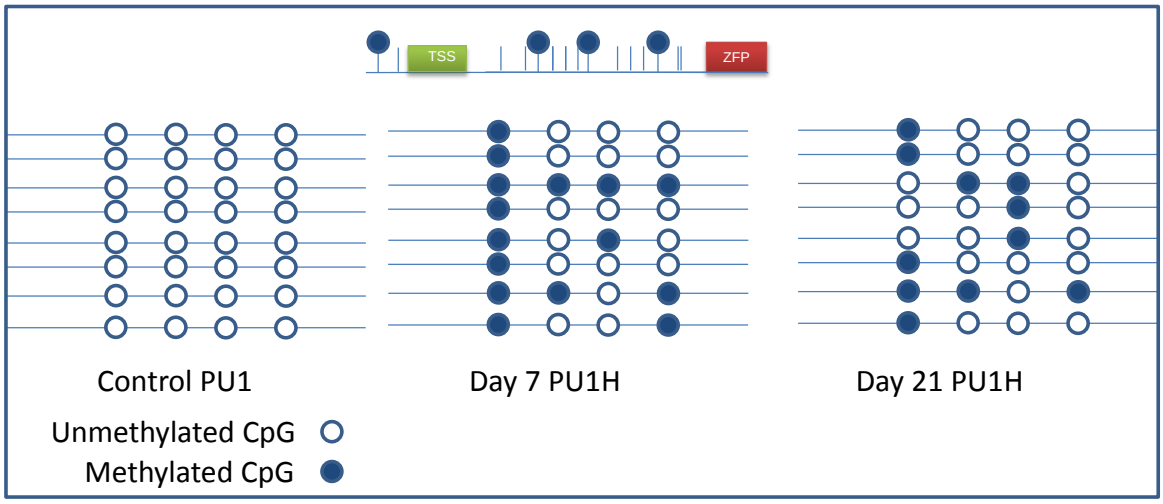


Figure 4-15 Clonal analysis of DNA methylation at HpaII sites.

PCR amplicons generated by a semi nested approach from bisulfite converted PU1H or PU1 targeting samples at Day 7 and Day 21 were subcloned into TOPO® vector. Eight selected colonies were assessed for methylation by sequencing. Open circles represent unmethylated HpaII CCGG sites, whereas closed circles show methylated CC^{me}GG sites.

To eliminate the possibility that the observed methylation patterns seen at longer time points were due to the effect of the plasmid expressing the methyltransferase not being diluted out *via* replication following transient transfections, PCR was performed to detect the presence of M.HpaII transcripts. RNA was extracted from the samples (1x10^6 cells) collected at time points Day 7, 14 and 21, followed by cDNA preparation as described previously. Primers were designed to amplify a 220bp region of the M.HpaII prokaryotic gene. HpaII transcription was evaluated at Day 7, 14 and 21. PCR amplicon samples were collected starting at cycle 20 at four cycle intervals up to 40 cycles. Negative controls used in the assay were PU1 and PU1Hh. Shown in Figure 4-16 is the result of the analysis. It is apparent that the plasmid bearing HpaII construct present at Day 7 is gradually diluted out and there is no HpaII detected at Day 21 samples. This clearly indicates that the observed methylation patterns are a direct consequence of the targeted methylation and furthermore that such patterns are epigenetically maintained. Previous similar analysis by our lab confirms loss of such vectors at about Day 12 [342].

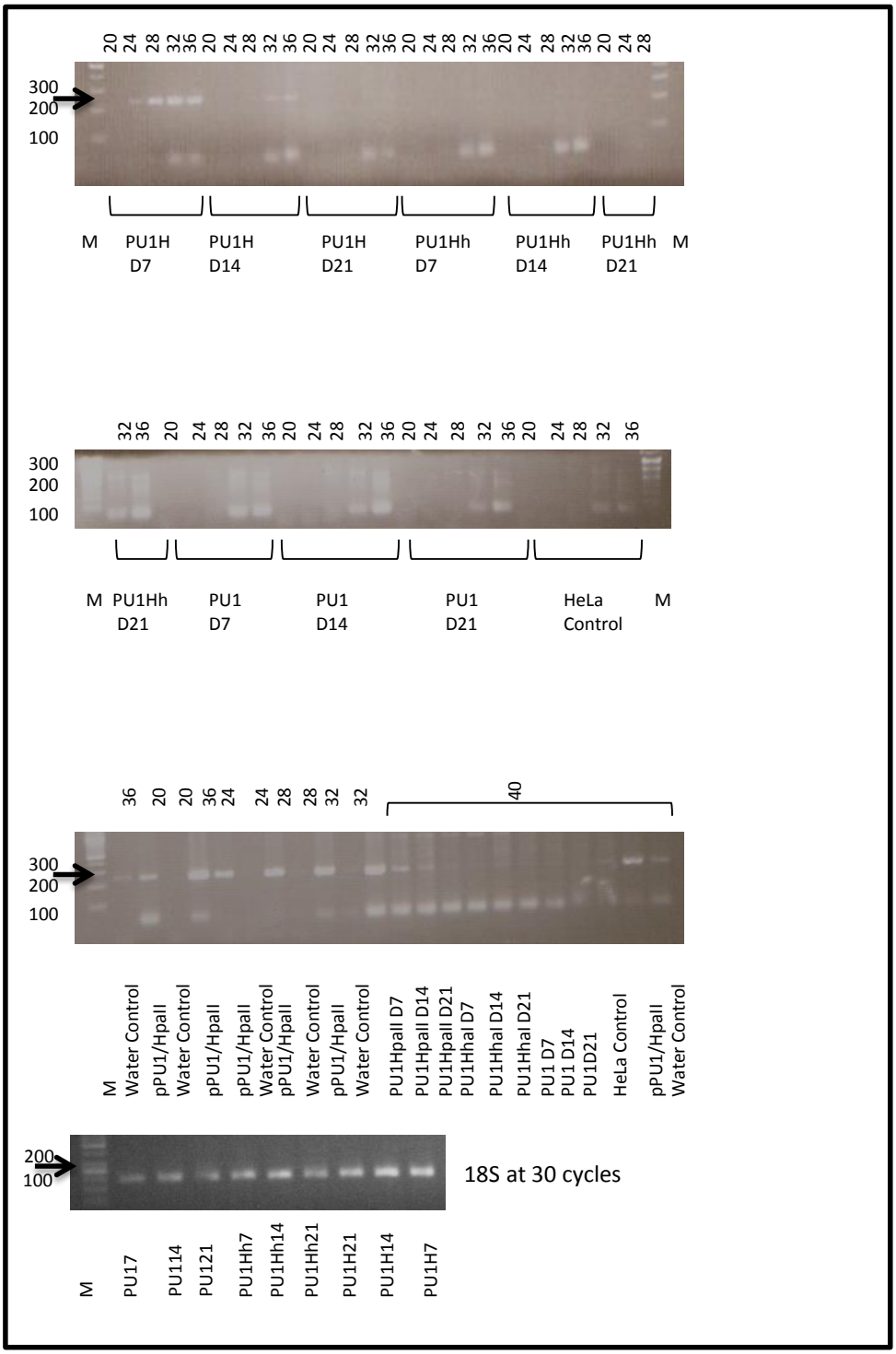


Figure 4-16 HpaII methyltransferase PCR

RNA was isolated and cDNA generated from $1x10^6$ of HeLa cells transfected with PU1H, PU1 and PU1Hh at the time points Day 7, 14 and 21. HpaII specific primers were used to amplify a 220bp region. PCR samples were collected starting at cycle 20 at four cycle intervals up to 40 cycles. A 220bp band observed at Day 7 for PU1H is not visible at Day 21.

4.9.1.2 *De novo* methylation spreads into the distal region of the *CDKN2B* following targeting

To investigate the mechanisms of *de novo* methylation spreading following targeted methylation delivery, a region in closer proximity to the ZF binding site was evaluated. Bisulfite converted samples of PU1H and PU1 at day 14 were PCR amplified using a new primer set to generate an amplicon 577bp in size covering a region upstream and downstream of ZF target site. Although the region downstream from ZF site contains no HpaII sites there are five CpG sites that could potentially be methylated via the methylation spread mechanisms.

There were four potential HpaII sites at this region for targeted methylation deposition but no methylation was observed. The region evaluated however contained 35 non HpaII CpG sites, 10 of which were methylated. This means that, there was a ~29% increase in regional DNA methylation levels at CpG sites that could be observed 14 days following methylation targeting as shown in Figure 4-17. While the physical link between the 5' and 3' region methylation pattern can't be categorically proven, the fact that all clones examined in the 5' region had HpaII mediated methylation and that a high percentage of 3' region clones are likely to share the methylation pattern shown in Figure 4-17, given cytosine peak intensities, suggests it is likely that 5' and 3' patterns are likely to be coincident on individual clones. Attempts to generate a larger PCR product encompassing both 5' and 3' regions failed to work, due to the size limitations associated with amplifying from bisulfite converted DNA.

Analysis of methylation distribution shows a uniform spread within the region. Interestingly, the region downstream of the ZF target site, where five CpG sites were available for methylation spread remained unmethylated, and may reflect the orientation of the methyltransferase with regards to the ZF binding site (i.e. is it pointing in the 5' direction from the ZF). It appears that *de novo* methylation patterns spreads within the 3' region of the promoter, whilst no spreading is observed in the 5' region relative to TSS. Whether this changes at much longer time-points is currently being assessed.

Figure 4-17 Methylation Spread Analysis – Bisulfite Sequencing

PCR products obtained from a library samples at Day 14 following targeting with ACMVPU1H, and ACMVPU1 constructs were gel cleaned as described in Chapter 2 Methods Section 2.8.2. and sequenced. Four HpaII sites highlighted in green are unmethylated. CpG sites upstream and downstream of ZF target site (blue) available for methylation and methylated are highlighted in black and red respectively. *R.BsmI* recognition sequence (GAATGCG) used for restriction analysis is circled.

Following demonstration of successful low density targeted DNA methylation, the effect targeting had on translation of the $p15^{INK4B}$ protein and transcription level of *CDKN2B* was examined.

To assess if any difference could be observed at the translational level samples were prepared for western blot analysis of $p15^{INKB}$ protein expression at time point Day 4, 14 and 21. Samples ($1x10^6$) were prepared for western blot analysis as described previously. Primary mouse monoclonal antibody, $p15^{INK4B}$ was used at 1:250 dilution in 3% (w/v) BSA. PU1H and PU1 transfection samples were used for comparison of protein expression levels shown in Figure 4-18. No changes in the protein levels were observed.

To look at the transcriptional levels of *CDKN2B* following methylation of the promoter, samples taken at Days 7,14 and 21 post transfection were examined using, RT-qPCR analysis by Taqman®. The method described in Methods Chapter 2 Section 2.13.1 was followed throughout. FAM labelled probe was used for the detection of *CDKN2B* transcription alongside the VIC labelled probe for 18S which was used as an endogenous reference in the assay Figure 4-18.

SDS – Sequence Detection System software was used to generate amplification plot and determine Cycle threshold points (Ct) for each sample The real time system uses fluorescent resonance energy transfer (FRET) generated by fluorogenic dual labelled probes consisting of a reporter and a quencher dye and the 5' nuclease activity of Taq DNA polymerase to amplify PCR products exponentially. For *CDKN2B,* a detection probe labelled with 6-carboxyfluorescein (FAM) and a carboxy-tetramethyl-rhodamine (TAMRA) quencher and primer were designed to produce a 169bp fragment spanning exons 1 and 2, to avoid potential amplification of highly homologous p16 exon 2 [411]. The 189bp 18S reference PCR was generated using Abi Human 18S VIC labelled control kit.

The ΔΔCt approximation method was used to evaluate relative expression of *CDKN2B* normalised to the endogenous housekeeping reference 18S. The primers were checked for efficiency by performing 10-fold dilutions and comparing the Ct values. Data analysis however revealed no significant changes in *CDKN2B* transcription levels.

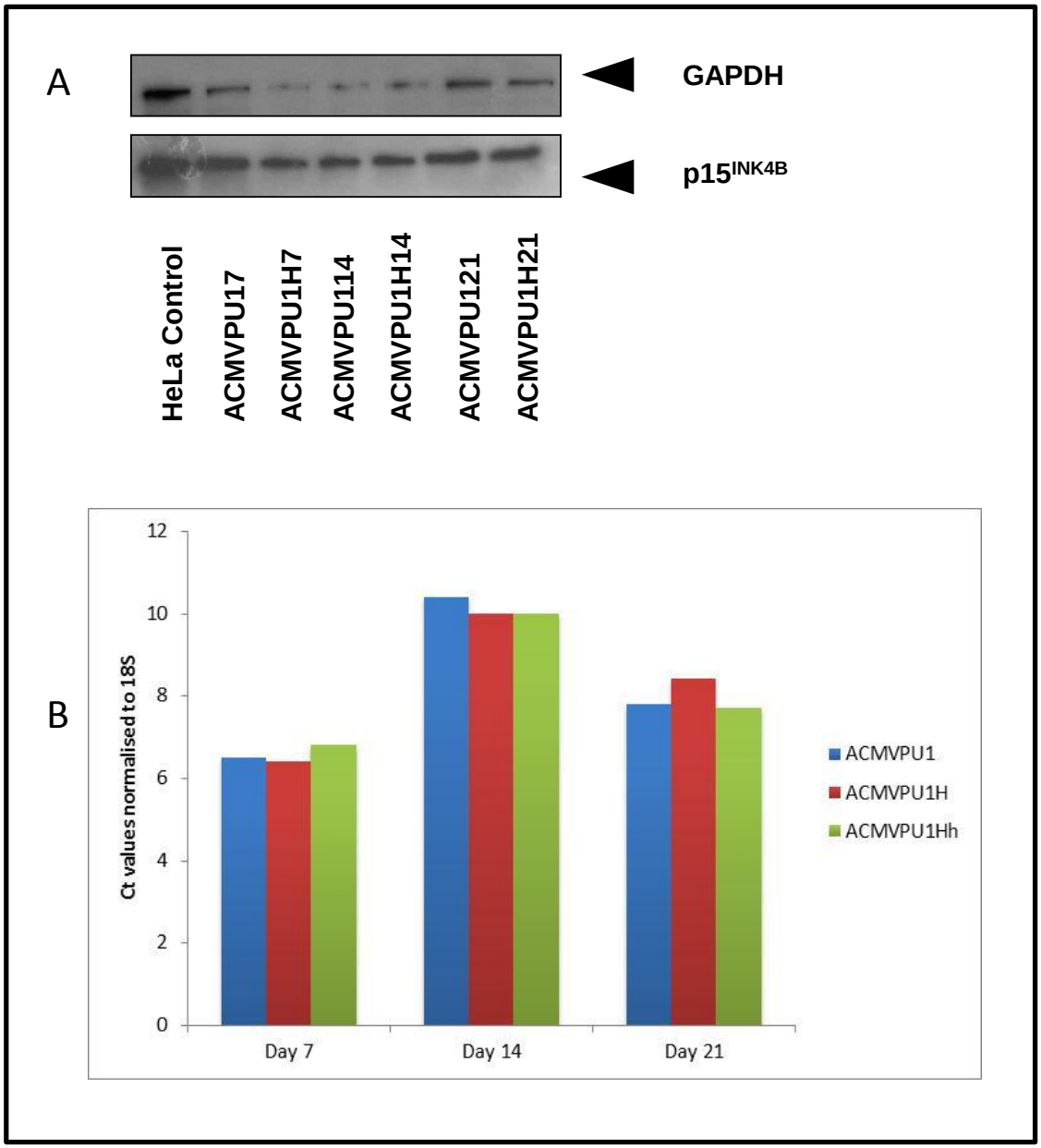


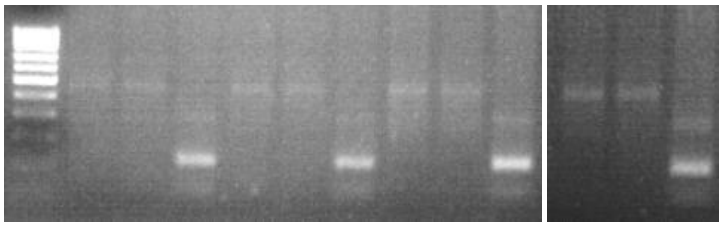
Figure 4-18 *CDKN2B* gene expression evaluation following targeted delivery.

HeLa cells were transfected with PU1H, PU1Hh, and PU1 and the samples ($1x10^6$ cells) were collected at time points Day 7 and 14 for RT-qPCR Taqman® analysis and Day 7, 14 and 21 for western blot analysis of *CDKN2B* transcription levels and $p15^{INK4B}$ protein expression. **A** Total cell lysate obtained from the transfected samples probed for $p15^{INK4B}$ and GAPDH which was used as a loading control. Cellular protein content was determined by BCA assay and the equivalent amounts of cell lysates were loaded per well. Arrowed are the observed bands corresponding to 15kD for $p15^{INK4B}$ and 37kD for GAPDH. **B** RNA extracted from samples transfected was used as a template to prepare cDNA which was then used as a template for *CDKN2B* relative expression analysis normalised to endogenous housekeeping gene reference 18S. There is a slight increase in Ct values observed at Day 7 samples.

4.9.9.1 High density methylation at *CDKN2B*

In an attempt to lay down higher density methylation in the *CDKN2B* promoter region ZFPs fused catalytic domains of eukaryotic DNMT3A, DNMT3B expressed either alone or in conjunction with DNMT3L as well as prokaryotic M.SssI methyltransferase were used. These methyltransferase target C residues within a CpG context and as already discussed in Figure 4-11 of this Chapter, there are 64 potential sites that could be targeted within *CDKN2B* promoter CGI using these enzymes. Having successfully targeted low density methylation to the promoter region, we aimed to investigate whether it was possible to impose higher methylation density and what the outcome of such targeting would be. In our studies we have used all the ZFP constructs (listed in Table 4-2). Additionally, ACMVPU1/3A, ACMVPU1/3B alone or co-targeted with ACMV/DNMT3L and ACMVPU1/SssI were used. The rationale for including DNMT3L into the study was based on earlier observations that this enzyme stimulates the DNA methylation activity of DNMT3A and DNMT3B through a direct interaction with its carboxyl terminal domain [412].

HeLa cells were transfected with a variety of constructs. Samples of transfected cells were collected at Day 4,7,14 and 21, and genomic DNA was prepared and bisulfite converted as described previously. The same primer sets that were used for COBRA analysis and bisulfite sequencing for assessment of M.HpaII (CCGG) and M.HhaI (CGCG) target sites were also used for this analysis. This allowed for a region containing 35 CG sites to be analysed for methylation status. Once generated, PCR products were restriction digested using *R.BsmI* and *R.Tsp5091* (for conversion efficiency confirmation) for ACMVPU1/3A targeted samples, but no methylation could be observed by this method. Bisulfite libraries were also sequenced. Figure 4-19 shows the restriction digest results for time points Day 7 and 14 for ACMVPU1A. Digests were also carried out at time points Day 4 and Day 21 with the same outcome (data not shown).



M 1 2 3 4 5 6 7 8 9 10 11 12

L1-ACMVPU1A D7

L2-ACMVPU1A D7 *R.BsmI*

L3-ACMVPU1A D7 *R.Tsp5091*

L4-ACMVPU1 D7

L5-ACMVPU1 D7 *R.BsmI*

L6-ACMVPU1 D7 *R.Tsp5091*

L7-ACMVPU1A D14

L8-ACMVPU1A D14 *R.BsmI*

L9-ACMVPU1A D14 *R.Tsp5091*

L10-ACMVPU1 D14

L11-ACMVPU1 D14 *R.BsmI*

L12-ACMVPU1 D14 *R.Tsp5091*

Figure 4-19 Methylation Analysis – Restriction digests

Restriction digests of PCR products obtained from samples following targeting with ACMVPU1A, and ACMVPU1 constructs. To ascertain whether methylation occurred using ACMVPU13A fusion, prepared PCR products were subjected to *R.BsmI* restriction. The enzyme recognizes and cleaves only when the site derives from methylated DNA, GCGGAATGCGCAGG to give two distinct bands 413bp and 59bp in size. To confirm bisulfite conversion has taken place *R.Tsp5091* was used as a conversion control check. The recognition sequence for this enzyme is AATT and is present at positions 34 and 168 of the amplicon only in the bisulfite converted DNA. Therefore, bisulfite converted DNA would yield two bands 134bp and 338bp in size only if successfully bisulfite converted.

No methylation could be observed using COBRA analysis so the prepared bisulfite libraries were sequenced as described in the previous section. The results of one of sequencing reactions, ACMVPU1A Day 4 is shown in Figure 4-20. Possible CG sites (35) that are available for methylation are underlined. Out of 35 possible sites only two were potentially methylated and were in the region where the bisulfite conversion efficiency was not 100% which means that the CGs at those sites are most likely the result of incomplete conversion. Similar lack of methylation was observed for all other targeted samples that were analysed: ACMVPU1A and ACMVPU1B coupled with or without DNMT3L expression at time points Day 4 and 14.

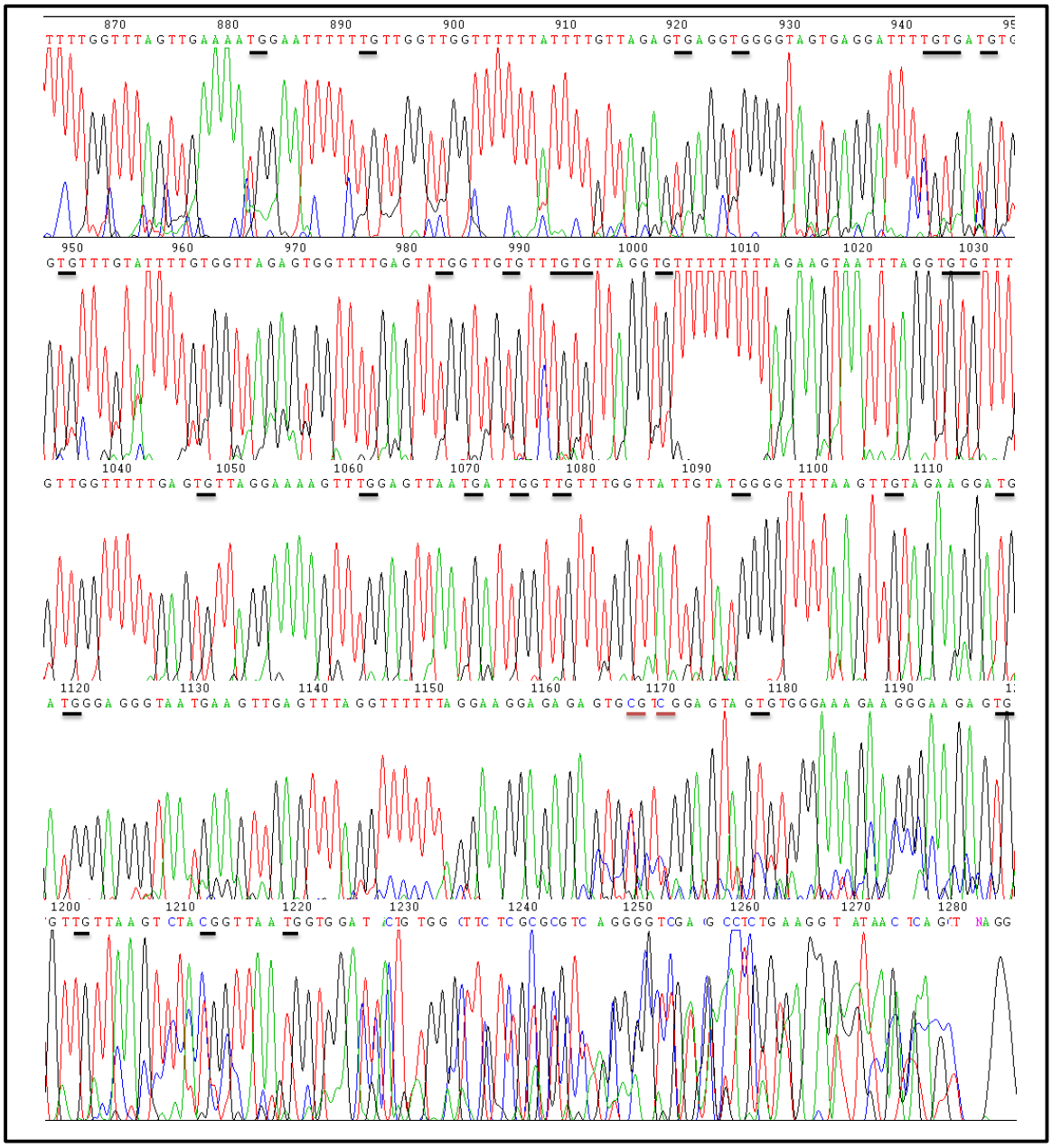


Figure 4-20 Methylation Analysis – Bisulfite Sequencing

PCR products obtained from samples at Day 4 following targeting with ACMVPU1A were gel cleaned as described in Chapter 2 Methods Section 2.8.2. No methylation was observed by COBRA (shown in Figure 4-16). ACMVPU1A transfected samples, prepared in the same way as for COBRA analysis were bisulfite sequenced commercially. The underlined sequence within *CDKN2B* highlighted in black shows all CG sites available for methylation. There were 35 CG target sites available in the region upstream from the PU1 ZFP binding site within the PCR region amplified and only two of them possibly methylated (highlighted in red).

4.10 Conclusion

For this project ZFPs have been expressed as fusions with various methyltransferases to target low and high density methylation to the *CDKN2B* promoter region. To the best of my knowledge no one has successfully exogenously imposed DNA methylation in the promoter region of an actively transcribing gene before in this manner. It has long been speculated that the open chromatin established by an active promoter actively excludes methylation from the region. Therefore, it was uncertain whether accessing an actively transcribing endogenous gene promoter region would be an issue or not.

Here it is demonstrated for the first time that indeed it is possible to methylate the promoter of an endogenous gene. *CDKN2B* promoter methylation is observed at Day 4 following transfections with ACMVPU1/HpaII, and ACMVPU1/HhaI. Initial library sequence analysis revealed that all of the HpaII target sites evaluated were methylated to varying extents. It is noteworthy that these methylated sites were in closest vicinity to the TSS, spanning the region between – 7 to +249 relative to TSS. Additionally, methylated HpaII sites (for the region analysed) were observed at 300 bp and as far as 579 bp downstream of ZF target site. This is much further than the $(Gly_4Ser)_3$ linker could feasibly reach, and suggests either DNA conformation effects, bringing the site closer to the ZF region, or diffusion of ZF/methyltransferase from its binding site.

Furthermore, *de novo* methyl marks once established could be detected up to Day 21 for M.HpaII constructs and up to Day 14 for M.HhaI. Day 21 was taken as the final time point for this project. Studies to look at the longer term methylation status of the promoter are on-going. The heritability of the newly established methylation patterns was further confirmed using restriction digest analysis and a finer resolution analysis of selected individual library clones.

Next, the observation of a distinct pattern of methylation spread to the 3' promoter region is an additional unique finding of this project. There is no obvious spreading of *de novo* methylation to the most immediate surroundings of the HpaII methylated sites. It could be speculated that the spread of methylation further into the 5' region of the promoter (which for the *CDKN2B* gene is a region harbouring the TSS) is inhibited by specific DNA sequences such as transcription

factor binding sites. However, a very distinct increase in methylation of CpG sites in the 3' region of the promoter nearer to the ZF target site has been observed. At Day 14 there is around 29% level of methylation in this region of the promoter. Surprisingly, there was no targeted methylation at the four available HpaII sites in this region. One possible explanation for this could be that ZF/methyltransferases are prevented from accessing the HpaII sites in the region of immediate vicinity due to active blocking by nucleosomal presence. Indeed the periodicity of the observed CpG methylation pattern is similar to the nucleosomal repeat unit of around 200bp. Figure 4-18 proposes a mechanistic model for targeted *de novo* methylation and subsequent spread into the distal CpGs of the promoter.

The PU ZF target site was initially selected because its recognition sequence was on the boundary region of the gene promoter. It could be speculated that this is a region marked by nucleosomal positioning. The "statistical positioning" model first proposed by Kornberg and supported by experimental evidence posits nucleosomes at the boundary regions between nucleosome free regions (NFRs) found at active or poised genes promoters [413]. The fact that there is an observed increase in methylation in this region further supports the possibility of nucleosomal occupancy. Global DNA methylation analysis at a single base resolution shows that nucleosome bound DNA is more highly methylated than flanking sequences suggesting that DNMTs preferentially target nucleosomal DNA [414]. DNMT3 enzymes were shown to act in heteromeric complexes with two methyltransferase active sites, creating a spacing to accommodate 10 nucleotides of DNA [415]. By what mechanism upstream region methylation might be "sensed" and what then invites CpG methylation to the nucleosome is unclear. Recently, a novel mechanism for DNMT3A/3B activity has been put forward which suggests that the levels of 5meC determine the bound levels of DNMT3A/3B. Furthermore, methylated DNA bound nucleosomes stabilize *de novo* DNMT3A/3B allowing for very little of it to remain free in the nucleus. Upon a reduction in DNA methylation, nucleosomal binding is disrupted leading to degradation of the unstable free DNMT enzyme thereby preventing aberrant *de novo* methylation [416] .

The observed epigenetic inheritance of the *de novo* DNA methylation pattern in the upstream region of targeted *CDKN2B* confirms DNMT1 maintenance activity presumably via UHRF1 which slides along the DNA and detects hemi methylated CpGs via its SRA domain by a base flipping mechanism [417] . Another part of UHRF1 may be operating at the histone level and signals to another interacting partners such as E2, an ubiquitin conjugating enzyme, *via* its RING domain, HDAC via SRA domain and G9a via PHD and TTD domain, to act on histone H3K9 [418] . However it was recently demonstrated that HP1 and UHRF1, alongside a host of other key regulators such as PCNA, Me-CpG binding protein 2, EZH2, and HDAC1 were not required for guiding DNMT3A/3B to the nucleosome [419] suggesting alternative localization mechanisms exists.

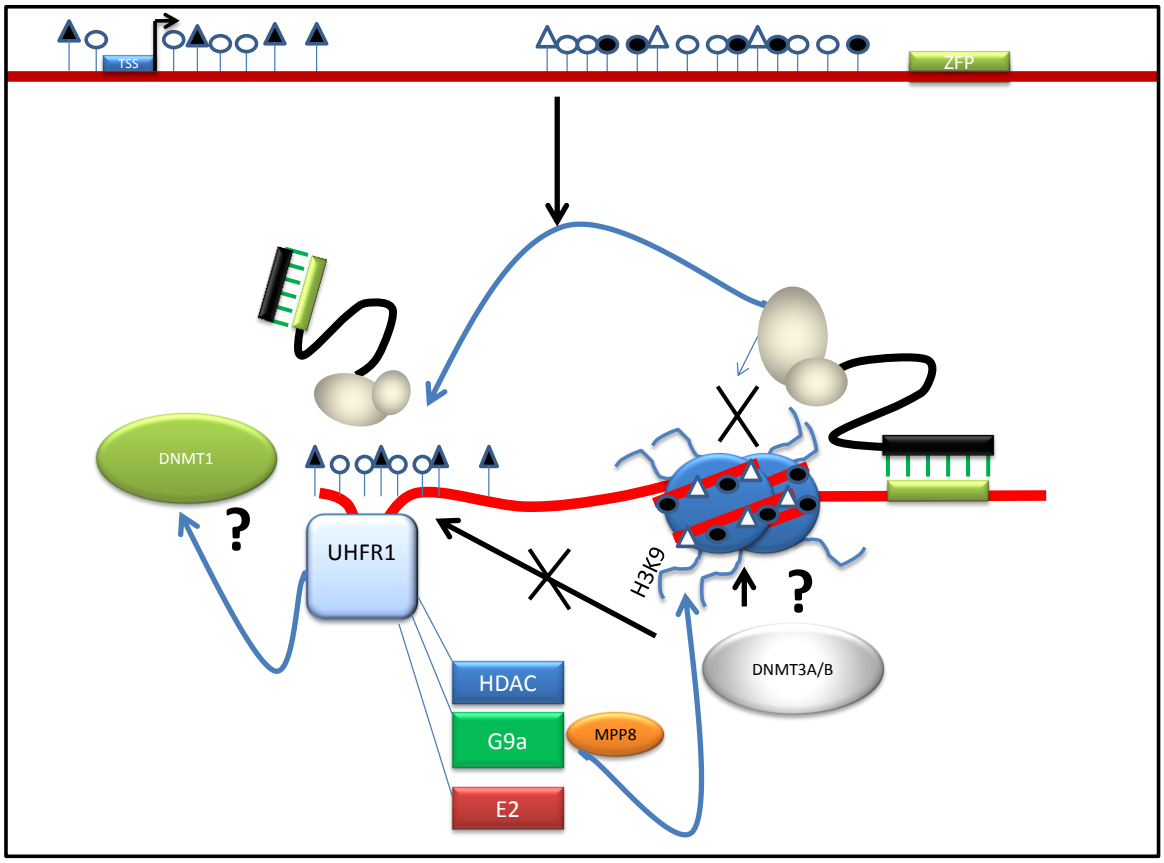


Figure 4-18 Schematic of a proposed model for the methylation distribution

Nucleosomal positioning adjacent to the ZF target site prevents access to the ZF/Mtase fusion to methylate HpaII sites. HpaII sites located approximately 300bp towards TSS were methylated. The observed maintenance of methylation indicates recruitment of DNMT1s probably aided by the activity of a "reader" UHFR1. This multidomain protein could provide a functional link for the cross talk between DNA methylation and histone modifications. The DNMT3A/B enzymes are known to selectively anchor to nucleosomes but the precise recruitment mechanism for this selectivity has not been fully determined. One possible molecular link recently put forward is the observed G9A DNMT3 co- operativity via MPP8 [420]. Another possible molecular mechanism might involve UHRF1 RING domain interaction with E2 an ubiquitin-conjugating enzyme. Histone ubiquitylation may provide an additional link in the recruitment of DNMT3s as recently demonstrated in ubiquitin-like proteins such as NEDD8 association with DNMT3B [421]. Additionally, an already established connection between H2A ubiquitination and DNMT3B could be another mechanism by which the enzyme is preferentially recruited to the nucleosome [422].

A progressive increase in DNA methylation patterns arising from discrete regions and spreading to adjacent sites has been described for a number of neoplastic transformations. For instance, methylation in three discrete regions of p16 CGI in human mammary epithelial breast cancer (HMEC) cells escaping mortality stage 0 (M_0) spreads into adjacent sites following continued cell passage leading to a gradual increase in methylation density across the island [423]. Similarly, for the *CDKN2B* promoter CGI, methylation density increases were observed in advancing MDS disease subtypes [424]. Finally, recent analysis of methylation expansion phenotype in CGI ,methylator phenotype (CIMP) showed hypermethylation spread in the 3' region in CIMP tumours [425]. The same study proposed a mechanism for the unidirectional expansion of methylation which occurred in 86.7% of subsets analysed. This mechanism postulates the involvement and targeted direction of methylation by DNA binding factors such as Sp1 and TFAP2A, but also some less validated candidates such as ZBTB7B and CNOT3. These findings add an additional dimension to an already complex model of methylation spread and regulation.

Assessment of the transcriptional regulation at the *CDKN2B gene* showed no change in the gene expression levels as a result of targeted methylation at the promoter region, either at the transcriptional or translational level. Although, this is a short – term study, so far the data suggest the low density methylation observed is insufficient to have an effect on gene expression level and suggest lack of formation of repressive chromatin structure at the *CDKN2B* promoter in response to targeting *de novo* methylation.

This is perhaps unsurprising as, evidence supporting the model of gene transcriptional silencing preceding the methylation has been reported. Kinetic studies to determine the timing of the epigenetic events involved in gene silencing using a stably integrated transgenic region in a chicken erythroid cell line showed rapid chromatin influenced silencing, associated primarily with H3 and H4 hypoacetylation and H3K4 methylation loss. These histone modifications were primary events in the transgene silencing, followed by the secondary events of H3K9 and promoter DNA methylation [426].

Evidence from *in vitro* studies suggests that promoter DNA methylation can influence gene expression in a process mediated by MeCP-1, and is dependent on methyl CpG density and

promoter strength. Densely methylated genes interact strongly with MeCP-1, and are frequently associated with gene repression. For sparsely methylated promoters the level of MeCP-1 interaction appears weaker, and can be disrupted according to the strength of the promoter, thereby allowing transcription of methylated genes [427].

During this project, I was not successful in delivering high density methylation using the catalytic domains of the DNMT3A/3B methyltransferases. The exact reasons are unclear, particularly given the fact that the DNMT3L component was simultaneously co expressed to enhance the enzymatic activity of the targeted DNMT3A/3B methyltransferases. The rationale behind co – targeting with the catalytically inactive protein DNMT3L is due to its demonstrated stimulatory effect on the catalytic activity of DNMT3A and DNMT3B [428]. Additionally, DNMT3L in heterodimer formation with DNMT3A and DNMT3B guides CpG site recognition in a periodic pattern [429] and attenuates formation of more uniform methylation patterns [430]. However, DNMT3L is expressed in ES but not in somatic cells [431] which may suggest alternative mechanism requirement for DNMT3 activity in these cells.

The lack of observed methylation at the target region can't be explained by ZF binding interactions given the fact that the same ZFs were successfully used for M.HpaII and M.HhaI targeting. It is possible that the ZF/Mtase component folding is different in this case and that could contribute to a failure to methylate. The low enzymatic activity associated with DNMT3A/3B catalytic domain Mtases is also unlikely to be the cause given the fact that the same catalytic domain DNMT3A and DNMT3B were previously demonstrated to have enzymatic activity when expressed as ZF fusions *in vivo* [432]. The catalytic domain of DNMT3A/3B fused with an engineered Cys2His2 zinc finger domain were previously successfully used to target three different artificial reporter systems; the TK (Humanherpesvirus 1 thymidylate kinase) promoter with added binding sites for GAL4 (UAS), the human c-Ha-ras gene promoter with added UAS and the immediate early IE175kpromoter of the Herpes Simplex Virus type 1 (HSV-1). The study demonstrated site specific methylation associated with the repression of the reporter genes.

What follows is an evaluation of several possible explanations for the observed lack of methylation targeting. Firstly, the above studies were based on plasmid episomal target vectors. There are probably differences in epigenetic regulation and nucleosomal structure of

the target episome compared to the *in vivo* setting of an actively transcribing endogenous gene. Secondly, their study used Human Embryonic Kidney 293T cells. The cell line was originally transformed by expression of the large T antigen from SV40 virus that inactivates pRb, leading to inhibition of cell cycle. The cellular epigenetic regulatory mechanisms between these cells and HeLa, which were used for this project may differ.

Thirdly, it is noteworthy that the same study reported a low level repression (~10%) using catalytically inactive DNMT3A suggesting a possible repressor recruitment by the enzyme.

Another but related point to the above is the possibility that the catalytic domain of the enzymes interacts with other protein components such as HDAC or HMTs which then "piggy-back" it away, affecting the enzyme's cellular localization. While evidence from previous studies points to predominant DNMT3 interactions mostly via their N-terminal domain, evidence for interaction with the catalytic domain of DNMT3B and HP1α and Suv39H1 also exist [433]. Additionally, the repair enzyme thymine DNA glycosylase (TDG) was shown *in vitro* to inhibit the methylation activity of DNMT3A [434] suggesting a DNMT3A C-terminus interaction.

Finally DNMT3A enzymes show a preference for CpG sites but are also able to methylate CpT and CpA, which are modified only twofold slower than CpG sites [435]. That would suggest that background levels of C^{me} seen in Figure 4-17 may not be due to incomplete bisulfite conversion but actually a consequence of genuine DNMT3A non-CpG methylation activity. Given that the COBRA restriction digest approach is hampered by poor site representation and bisulfite sequencing can't easily distinguish between methylated and unconverted samples, a different approach is ultimately required to assess the methylation pattern of these samples.

Attempts to target selective high density CpG methylation using the prokaryotic M.SssI methyltransferase were unsuccessful. This enzyme was previously demonstrated to target methylation when fused to zinc finger proteins, in yeast [436]. M.SssI is frequently used as an enzyme of choice to methylate DNA templates for *in vitro* studies on the role of DNA methylation. A large body of studies involving M.SssI include evaluation of promoter methylation mediated transcriptional control of viral and mammalian genes [437-439], demethylation studies [440], and investigations on effects of folate deficiency on DNA methylation [441].

One very interesting feature of this enzyme is its topoisomerase activity. It has also been shown that methylation by M.SssI in the presence of Mg^{2+} becomes distributive rather than processive resulting from the enzyme's decreased affinity for DNA. These two enzymatic functions are not interdependent, with either one or the other activity being detectable [442].

Following HeLa transfection with ZFP/M.SssI vector under the control of CMV promoter, no expression was detected by western blot. We assume that this is due to the M.SssI methylating the plasmid CMV promoter at high density, thereby rendering the expression cassette inactive. Evidence in support of this comes from studies showing strong and ubiquitously expressing promoters being susceptible to CpG methylation[443].

Attempts to clone M.SssI into different expression systems were previously shown to be unsuccessful, presumably due to toxicity issues associated with the expression of this gene in *E.Coli* [444]. The original M.SssI gene is known to have unusual codon usage, with four TGA codons presumed to code for tryptophan in the parent organism. Thus, issues of mammalian expression have to be considered in this context, coupled with the fact that M.SssI is a highly catalytic methyltransferase, with a very fast k_{cat} turnover rate[444].

Chapter 5

Investigating the role of cellular factors in maintaining the *CDKN2B* promoter in a methylation free state *in vivo*

5.1 Introduction

Using rationally designed ZF/HpaII methyltransferase fusion constructs I have demonstrated that it is possible to lay down site-specific methylation in the promoter region of the *CDKN2B* gene. Furthermore, I showed that this epigenetic tag is maintained through successive cell divisions and leads to the spread of methylation within the CGI in the 3' direction as discussed in the previous chapter.

Additionally, there appears to be no effect on *CDKN2B* gene transcriptional status as a result of this *de novo* methylation process. It could be speculated that promoter methylation driven gene silencing *in vivo* evolves from a series of events resulting from a substantial build-up of methylation rather than a short exposure low density event. Only once the region is "captured" by the spread of DNA methylation, can interference with transcription factor binding occur and stable gene silencing be established [445]. Studies on human (with CGI) and mouse α-globin and human γ-globin (without CGI) genes showed a methylation density requirement in transcriptional regulation. Low density methylation forms a weak complex with MeCP-1/NuRD which can be sufficient for repression of a weak promoter. However, when the promoter is strong, the complex is disrupted, allowing the transcription of a methylated gene [446].

So what are the exact mechanisms associated with the spread of methylation from a methylated "seed" foci? What keeps some promoter regions methylation free while the others succumb to methylation? These are still unknowns, despite our knowledge about the promoter regions of actively transcribing genes. In an attempt to answer these questions, it is essential to regard the promoter region of a gene as more than a simple transcriptional switch.

5.1. 1. Transcriptional Control at eukaryotic promoters

Gene promoter regions are sites at which the symphony of gene expression is orchestrated and where many epigenetic, as well as non-epigenetic instruments, play their part. Promoters contain a number of specific regulatory sequences such as transcription factor binding sites, transcription start site and TATA box recognition elements. Transcriptional regulation of *CDKN2B* has been reviewed in some details, in previous chapters.

An additional level of transcriptional control is provided by chromatin structure and nucleosomal organization in the promoter region as discussed in Chapter 1. Multi-protein complexes involved in the modulation of higher order chromatin structure helps establish transcriptional regulation. These complexes also mediate binding to both promoters and enhancers, ensuring cross talk between activators and repressors at transcriptional initiation sites, leading to specific nucleosomal positioning and new transcript generation. ATP provides energy for these complexes to ensure both non covalent chromatin signature modifications and acquisition or removal of histone tail marks. In principle, transcriptional repression can be achieved by the creation of compacted inaccessible chromatin, while transcriptional activation can be induced by creating open, accessible chromatin. However, these mechanisms of chromatin re-organization are in themselves not sufficient to regulate gene expression. Other factors such as transcriptional activator, repressor and mediator complexes and basal transcription factors are also needed.

Further, gene regulation is enabled by the activity of enhancers, silencers and silencing barrier insulators which sometimes act over long genomic distances, up to a mega base away from their target gene [447]. Silencers are DNA elements which play a role in gene insulation by compartmentalization of different chromatin domains. The nuclear periphery is enriched for heterochromatic silenced loci and the repressor proteins associated with it, which help stabilize transcriptional silencing.

Promoters are often positively and dynamically regulated by enhancers, which can act in a distance and orientation independent manner and aid repositioning of genes at transcription factories via chromatin loop interactions [448].

The activity of enhancers and silencers is tightly regulated through the effects of insulators which can interfere with promoter communication and act as either enhancer blockers or silencer barriers [449]. Insulators can also serve as heterochromatin barriers, to protect a gene from neighbourhood influences.

Additionally, insulators were shown to be capable of protecting promoters from DNA methylation as demonstrated for the multifunctional epigenetic regulator CCCTC-binding factor (CTCF) at the CGI of the pRb gene [450]. CTCF which has both insulating and chromatin barrier activity, was recently shown to co-localize at the INK4 locus, providing a barrier against heterochromatin spread [451]. Several CTCF binding sites were identified at this divergent region, and the binding shown to be DNA methylation sensitive [452] [453]. Furthermore, CTCF regulates the transcription of genes at this locus [453].

The cis- acting transcription factor - Sp1, which recognizes G - rich sequence motifs, was demonstrated to provide methylation protection at the CGI elements of both the mouse and hamster *APRT* gene [454, 455] with Sp1 deletion leading to CGI methylation. Sp1 *per se* is not required for the unmethylated state of this CGI, as it was shown that even in Sp1 null ES cells the *APRT* gene remains unmethylated [456]. Recent evidence from genome-scale methylation profiling of the human brain, where a number of putative zinc finger protein binding sites were found to be over-represented at the boundary regions of methylation-resistant CGIs, lends support to the existence of additional routes [457].

An example of a barrier element in vertebrates with a role in defending a gene from silencing by DNA methylation is Vascular Endothelial Zinc finger 1 (*VEZF1*). A study carried out on the chicken β-globin gene marked at the 5' region by an HS4 insulator, showed that this barrier element mediated gene promoter protection from DNA methylation-mediated silencing independent from CTCF –mediated enhancer blocking or USF mediated recruitment of active histone modifications [458].

An additional observation was that the *VEZF1* transcription factor could serve as a barrier/anti-methylation factor at the *APRT* gene promoter, and given that it recognizes the same G rich

sequences as Sp1 frequently observed at CGIs, it was postulated that VEZF1 may play a similar role in promoter DNA methylation protection. The role for VEZF1 as an epigenetic regulator is further supported by the contradictory observation of global, genome wide DNA methylation loss including LINE1 elements and minor satellite repeats in VEZF1-/- mouse ES cells which is also associated with substantial DNMT3B downregulation [459].

5.1.2. Active demethylation

An alternative mechanism for maintaining the methylation free status of gene promoters is that methylated CpGs are continuously removed through the activity of various DNA demethylase pathways. DNA methylation is known to be a dynamic process with active demethylation being observed during specific stages of development [460]. Furthermore, a cyclical pattern of DNA methylation has been observed at the transcriptionally active promoter of *pS2/TFF1* gene upon oestrogen activation. During this transcriptional cycling, a periodic, strand -specific methylation/demethylation is observed and is maintained by the dual activity of DNA methyltransferases which lay down CpG methylation and actively demethylate methylated CpGs through deamination [461]. There is a handful of examples of active methylation and demethylation of specific CpGs such as IL-2 promoter demethylation in somatic cells [462] or strand specific CpG methylation/demethylation [461, 463].

Another better established deamination enzyme is activation induced (cytidine) deaminase AID involved in the removal of the amino group from cytidine converting it into uridine, which is then recognized as a thymidine and subjected to DNA B.E.R. repair [464].

This, however, is only a part of the picture. The process of active demethylation is most likely context dependent and achieved *via* multiple mechanisms with a number of enzymes playing a role. One proposed enzyme was a member of the methyl CpG – binding domain family, MBD, related to the MeCP2 transcriptional repressor and reportedly involved in hydrolysis of 5-methylcytosine to cytosine and methanol [465, 466]. The reaction itself is, however, thermodynamically unfavourable due to the high activation energy required to break O-H and C-

C bonds to form C-H and C-O bonds, and now it seems more likely that this enzyme was a glycosylase.

Two other noteworthy demethylation pathways have been proposed. One involves the action of 5-methylcytosine DNA glycosylase to remove methylated cytosine and leave deoxyribose ready for local DNA repair mechanisms [467].

An alternative pathway involves the creation of an intermediate in the demethylation process by conversion of 5-methylcytosine to 5-hydroxymethylcytosine (5-hmeC) [468]. A group of enzymes proposed to be involved in this process and which could in theory be targeted to CGIs are Ten-eleven translocation TET family members; TET1, TET2 and TET2 are Fe^{2+} and α-ketoglutarate–dependent dioxygenases, which successively oxidise 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5-hmeC), 5-formylcytosine (5fC), and 5-carboxycytosine intermediates [469]. TET1 is found as a fusion partner of the *MLL* gene in AML which possesses a domain related to the CpG-binding CXXC domain found in *CFP1.* TET1 was shown to catalyze the conversion of 5mC to 5-hydroxymethylcytosine (5-hmeC) in HEK293 cells [470]. Recent studies have shown preferential localization of TET1 at CGIs in mouse ES cells and depletion of TET1 results in increased CpG methylation at CGIs [471, 472]. TET proteins were shown to co-operate with the base excision repair enzyme thymine-DNA glycosylase (TDG) suggesting a very important role for TET in epigenetic regulation [473].

It could be hypothesised that CGIs are continually subjected to sporadic *de novo* methylation, which is swept clean by a mechanism involving oxidation of 5-methylcytosine. Defects in such a mechanism such as TET2 mutations [474] could predispose the CGI to *de novo* methylation, as seen in many cancers. Recent evidence to support this idea comes from a study in mouse ES cells, where TET1 knockdown led to methylation of Nanog promoter and its downregulation [468].

5.2 Chapter Aim

On the basis of information gathered in the previous chapter regarding targeted methylation at the *CDKN2B* promoter, epigenetic inheritance of the *de novo* methylation patterns, and the effects on transcriptional regulation of the gene, it became apparent that promoter protection might be an issue warranting further investigations. Existing data indicate some spread of methylation following targeting but there is no evidence of gene repression, suggesting a protective mechanism exists in this region. In this chapter, I have begun investigating the relationship between two postulated mechanisms for promoter methylation protection; insulator and active DNA demethylation at the *CDKN2B* promoter region, with a view to their potential ultimate role in aberrant CGI methylation in cancer and leukaemia.

5.3 Results

5.3.1 *VEZF1* is enriched at the *CDKN2B* promoter

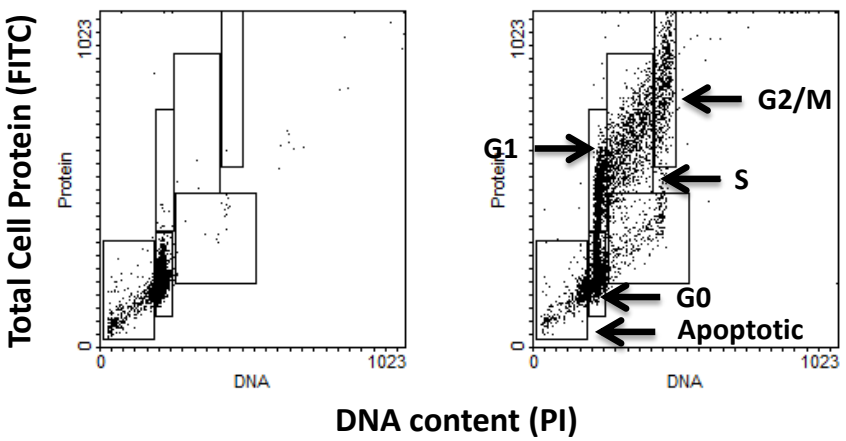
Using the GCG Wisconscin Package TF software, the proposed *CDKN2B* regulatory region was screened for putative TF binding sites. The results were then validated by further literature searches. Several Sp1 binding sites (GGACAGGGGGCGGAGCCTAAG), (AGAGCGAGGCGGGGCAGTG), (CTCTGCTCCGCCTACTGGG) and a putative *CTCF* site (CACCCCTCTCTTATGAT) were highlighted in the 5' region of the promoter.

Cell lysate samples of HeLa, Raji, K562, KG1 and HL60 cell lines which have well characterized *CDKN2B* expression and promoter methylation profiles were assessed for VEZF1 expression. Also, used were whole blood isolated primary T cells. The T cell isolations were kindly conducted by Cell Cycle group members (H. Milewicz, E. Hamilton and S.Orr). Non-activated human T cells

from healthy donors were isolated by negative selection from the peripheral blood. These quiescent T cells were then stimulated with CD3/CD28 magnetic beads (Dynal) for 48 hours.

Following magnetic bead removal by mechanical disaggregation, samples were taken for cell cycle analysis at G_0 – resting phase and stimulated to enter the G_1 cell cycle stage. This was done in order to confirm that T cells had responded to the stimulus and progressed through the $G_0 \rightarrow G_1$ commitment point [475], and to determine what proportion of T cells has entered the cell cycle (The procedure for T cell isolation, stimulation, Cell Cycle Analysis is described in Methods (Chapter 2, Sections 2.20 and 2.21).

The samples were fixed in 70% (v/v) ethanol and stained with FITC for total protein content and PI for DNA content and analysed by flow cytometry. At 48 hours, 20% of cells were in G_1 (Figure 5-1).



Cell Cycle Stage	Control	CD3/CD28 stimulated
G0	91.70	35.97
G1	0.14	19.84
S	0.05	17.63
G2/M	0.01	10.44
Apoptotic	0.24	5.69

Figure 5-1 Flow Cytometry Analysis of T Cells pre and post stimulation

Quiescent primary T cells were stimulated with CD3/CD28 beads for 48 hours. Cell samples were taken at 48 hours. Cells were fixed in cold 70% (v/v) ethanol and stained with PI (DNA content) and FITC (protein content). The percentage of cells in each cell cycle phase was determined by flow cytometry.

If VEZF1 has a role in promoter methylation protection as discussed in 5.1, as well as potential transcriptional contribution, it was necessary to look at the *VEZF1* expression status in various cell lines with known *CDKN2B* methylation status. Raji and KG1 have a partially methylated, whilst HL60 and HeLa have an unmethylated *CDKN2B* promoter. The K562 cell line has a deletion of the *CDKN2B* gene locus [476]. *VEZF1* expression was assessed by western blot analysis, and the results are shown in Figure 5-2 below. *VEZF1* expression was confirmed in all of the cell lines and primary cell samples with the exception of KG1 which showed no expression. KG1 and Raji although both exhibiting 50% methylation in the *CDKN2B* promoter region are cell lines of different lineage. Raji is B lymphocyte derived and KG1 is of human myeloid lineage.

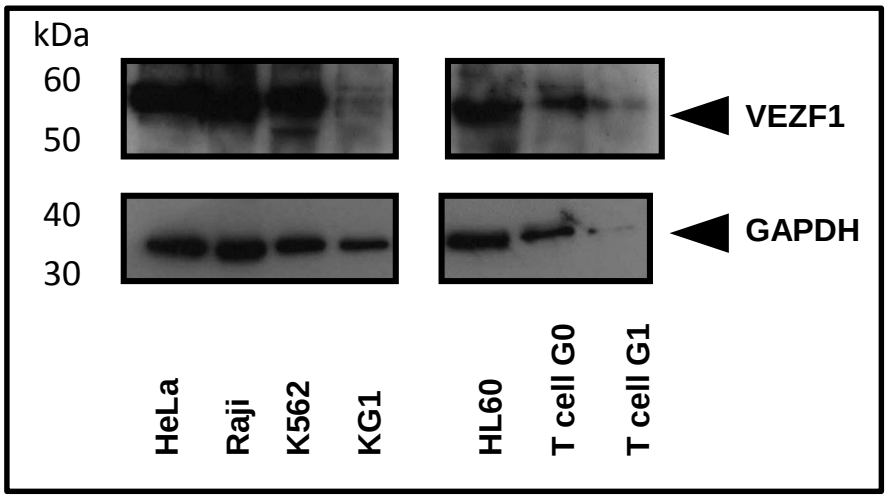


Figure 5-2 *VEZF1* protein expression in various cell lines and primary T cells.

VEZF1 protein expression was examined in HeLa, Raji, K562, KG1, HL60, quiescent and cell cycling primary T cells all at $1x10^6$, using rabbit polyclonal anti *VEZF1* antibody (sc-98278 X) at 200 µg/ml at 1:400 dilution in PBS 5% non - fat dry milk. *VEZF1* 56kDa protein expression was confirmed for the majority of cell lines, except KG1 which showed no expression of *VEZF1* protein. Anti-GAPDH goat polyclonal antibody was used at 1:10000 in PBS 3% BSA as a loading control to detect a band of approximately 36 kDa.

Having determined *VEZF1* expression in a majority of cell lines, I wanted to firstly assess whether *VEZF1* is recruited to the *CDKN2B* promoter. ChIP pull downs on HeLa cells were conducted using rabbit polyclonal anti *VEZF1* antibody (5µg) as described in Methods Chapter 2 Section 2.13.6. Negative control for the pull down was an IgG rabbit polyclonal antibody. PCR primers spanned the 5' region of the *CDKN2B* promoter (-890 to – 757 relative to the TSS) to generate a product 133 bp in size. A set of primers to detect the Endothelin 1 (ET-1) gene as a positive control was also designed. This gene was previously shown to engage a *VEZF1*/DB1-responsive element *via* a 6 bp motif ACCCCC located 47 bp upstream of the endothelin-1 transcription start site [477]. The result in Figure 5-3 shows significant *VEZF1* enrichment at the *CDKN2B* promoter. The result is representative of triplicate experiments. That VEZF1 binds to the *CDKN2B* promoter was hitherto unknown. GAPDH primers were used as an additional assay control. For input, 1/20th of the total amount was used per IP.

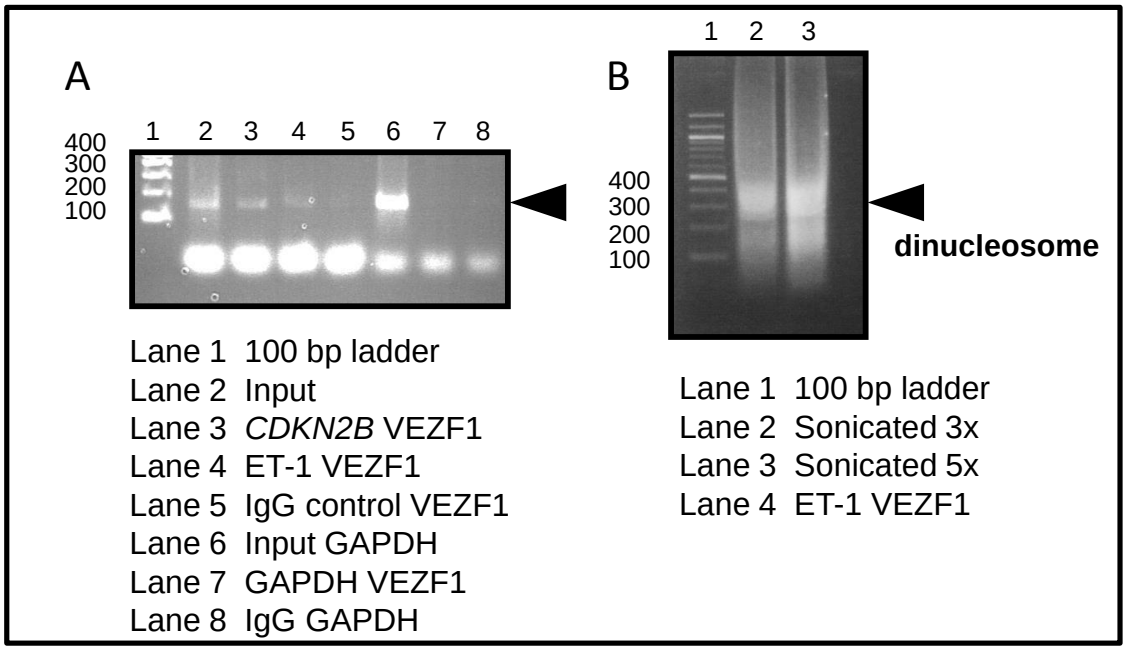


Figure 5-3 ChIP of *VEZF1* at the *CDKN2B* promoter

A HeLa cell ChIP using polyclonal rabbit anti-*VEZF1* antibody. Arrowed is a 133 bp PCR amplicon of the *CDKN2B* region and ET-1 region which was previously shown to engage *VEZF1*. IgG control in the last lane shows no band as expected. GAPDH was used as a negative control for each sample. **B** HeLa cells were x- linked and sonicated in a Bioruptor® (Diagenode), either 3 times or 5 times for 20 seconds on medium power setting.

.

Following confirmation of *VEZF1* binding at the *CDKN2B* promoter region, I asked what effect VEZF1 silencing would have on the promoter methylation pattern at the *CDKN2B* promoter, based on the hypothesis that it might mediate protection from methylation.

It was speculated that if VEZF1 has a role in *CDKN2B* promoter immunity to methylation then it would be interesting to address whether repeated knockdowns of VEZF1 would eventually lead to the *CDKN2B* promoter succumbing to more extensive *de novo* methylation and what its effect on *CDKN2B* transcription might be.

Repeated knockdowns of *VEZF1* were carried out using Silencer ® Select pre-designed siRNA (Ambion) at 4nM using Amaxa ® nucleofection at 4 day intervals. A positive control for transfection efficiency assessment was BLOCK-iT™ Fluorescent Oligo (Ambion, 2013). Transfection efficiency was assessed 48 hours following nucleofection using FACS, and for fluorescent control was shown to be ~98%. VEZF1 knockdown was monitored at 24, 48 and 96 hours following nucleofection and confirmed by western blot under the same conditions as previously described (Figure 5-2).

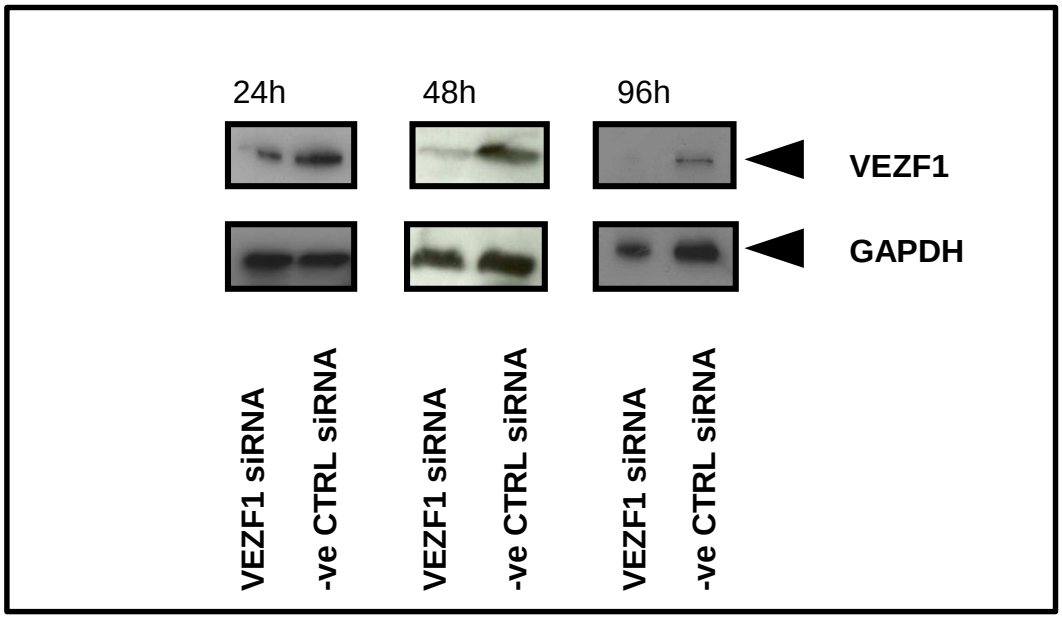


Figure 5-4 VEZF1 silencing in HeLa cells

Western Blot of VEZF1 siRNA knockdown in HeLa cells. HeLa cells were nucleofected with VEZF1 specific siRNA. Cells at $1x10^6$ were collected at 24, 48 and 96 hours following transfection and analysed by western blot using rabbit polyclonal anti *VEZF1* antibody (sc-98278 X) at 200 µg/ml at 1:400 dilution in PBS 5% non -fat dry milk. *VEZF1* 56kDa protein expression was unaffected by control siRNA, but a clear VEZF1 downregulation was observed for VEZF1 siRNA targeted sample. Anti-GAPDH goat polyclonal antibody was used at 1:10000 in PBS 3% BSA as a loading control to detect a band of approximately 36 kDa.

Repeatedly nucleofected samples were analysed for $p15^{INK4B}$ expression by western blot at four day intervals. This time point was selected to allow time for the cells to recover in terms of cell numbers following repeat transfections. Western blot analysis of $p15^{INK4B}$ shown in Figure 5-5 led to the interesting observation of an emergence of a potential p15 variant, 10kDa in size, following subsequent nucleofections, but not the first.

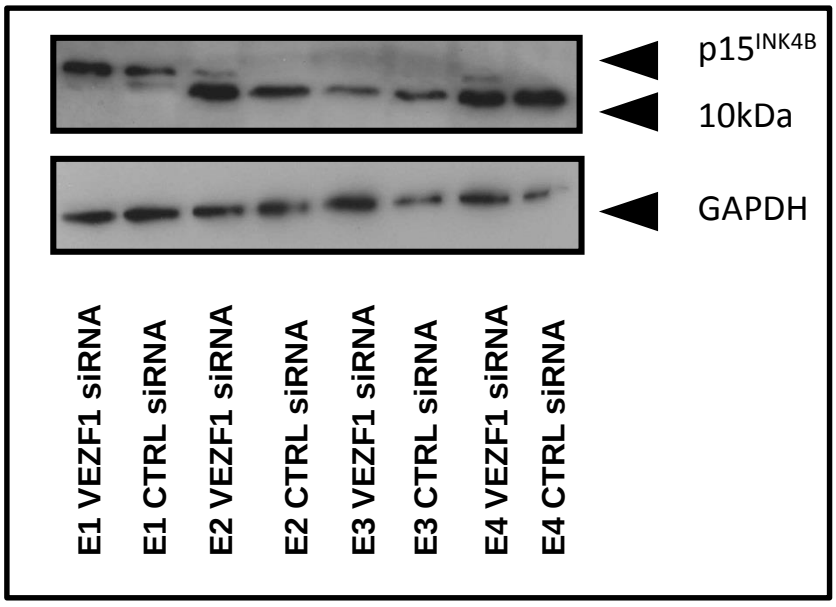


Figure 5-5 Alternative *CDKN2B* splice variant

Western Blot showing four experiments (E1 to E4) of repeated VEZF1 siRNA and a scrambled control siRNA nucleofections. $p15^{INK4B}$ expression was analysed as described previously in Chapter 4 Section 4.9.1.2. HeLa cells transfected with VEZF1 si RNA and a –ve control scrambled siRNA were assessed for $p15^{INK4B}$ expression at 4 day intervals. Result clearly shows a loss of $p15^{INK4B}$ expression and establishment of smaller size variant for both transfected samples following second nucleofection procedure.

Further investigations suggested two strong possibilities for the identity of this 10kDa band induced by repeated nucleofection. The *CDKN2B* gene has two coding exons encoding three transcript variants, p15, p15.5 and p10 which are functionally indistinguishable, with the main difference being that p15.5 is an N terminal extended form of p15 generated from an alternative inframe translation initiation codon [478]. The third variant p10 derives from alternative splicing within intron1. This protein is capable of cell cycle restriction, but it does not interact with CDK4 or CDK6 [479]. This splice variant has been shown to be induced by oncogenic insults and TGF-β treatment.

In mouse embryonic fibroblast (MEFs) and NIH 3T3 cells it was shown that p10 is able to induce cell cycle arrest and inhibit foci formation and anchorage-independent growth in a p53-dependent manner. This suggests that the *INK4B* locus with its different splicing variants can be involved in tumour suppression *via* the two key molecular routes – the p53 and retinoblastoma pathways [480].

In order to further ascertain that differential splicing of the *CDKN2B* gene results in the expression of p10 at the expense of p15, a set of p10 and $p15^{INK4B}$ specific primers spanning the region between the exon 1 and exon 1β/exon 2 boundary was designed [481]. However, there was no time left to conduct this experiment.

5.3.2 Global 5-hmeC signal intensity changes significantly during G_0-G_1 transition

I also wished to examine a potential role for TET family members in maintaining a methylation free status of the *CDKN2B* promoter. However, cultured cell lines exhibit very low levels of 5-hmeC. Evidence so far using an enzymatic approach of *in vitro* analysis of *R.MspI* and *R.HpaII* sensitivity to glucosylation of 5-hmeC showed that in HeLa and NIH3T3 5-hmeC amounts are nearly undetectable [482]. Another set of observations in HeLa cells indicates that the global levels of DNA methylation decrease during the G1 phase of the cell cycle and increase during the S phase of the cell cycle [483]. Therefore, attempts to ascertain the role of TETs in epigenetic regulation at the *CDKN2B* promoter in HeLa cells would be futile. As a prelude, to an analysis

of the effect of overexpression/knockdown of TETs on *CDKN2B* promoter methylation/demethylation I decided to examine TET and 5-hmeC levels in human primary T cells.

Relatively little is known about the changes in the epigenetic landscape associated with DNA 5-hmeC during cell cycle entry G_0-G_1. I wished to analyse levels of 5-hmeC during G_0-G_1 transition and investigate if there are changes associated with 5-hmeC levels and cell cycle entry. Total DNA 5-hmeC and meC content was compared in human T lymphocytes in a quiescent (G_0) and proliferating (G_1) state in response to an antigenic stimulus. T cells can remain in the peripheral blood for many years in the quiescent state before entering the cell cycle. Upon co-stimulation via CD3/CD28 cells commit to entry into the cell and growth cycle and expression of T cell effector functions [484].

Isolated quiescent and stimulated T lymphocytes were counted and analysed for cell cycle and purity as shown in Figure 5-1. Immunofluorescence staining was used to assess changes in 5-hmeC levels. T lymphocytes cell suspensions at $5x10^4$ were cytospun to adhere to poly-Lysine slides as described in Chapter 2 Methods, Section 2.20.2. The slides were prepared for immunofluorescence staining as described in Chapter 2 Methods Section 2.15. Briefly, cells were fixed in 4% paraformaldehyde for 15 minutes, washed with PBS and permeabilized with 0.4% Triton X-100 in PBS. Permeabilized cells were then denatured with 2N HCl for 15min and neutralized with 100mM Tris-HCl (pH8.5) for 10min before blocking for 1h with blocking buffer (10% goat serum). The cells were then incubated with polyclonal rabbit 5-hmeC specific primary antibody (Active Motif, 39769) at a 1:250 dilution in 10% goat serum in PBS overnight in a humidified chamber at 4 °C.

Following three consecutive 5-min washes with PBS, cells were incubated with secondary antibodies at 1:400 dilution in PBS. Two different secondary conjugated antibodies were used for the detection: Alexa Fluor ® 488 (rabbit polyclonal) and Alexa Fluor ® 647 (rabbit polyclonal). This was followed by DAPI staining for 15 minutes at room temperature. Finally, cells were washed again three times with PBS and mounted in the mounting medium. Images were acquired using Zeiss immunofluorescence microscope Axiovert 200 and Axiovision software.

We observed a significant increase in the 5-hmeC signal intensity during G_0-G_1 transition (shown In Figures 5-6 and 5-7). Figure 5-6 shows 5-hmeC signals at G_0 and G_1 cell cycle stages. Due to the high intensity of FITC signals at G_1 vs. G_0 an alternative secondary conjugated antibody was evaluated. This gave much clearer and more defined signals shown in Figure 5-7.

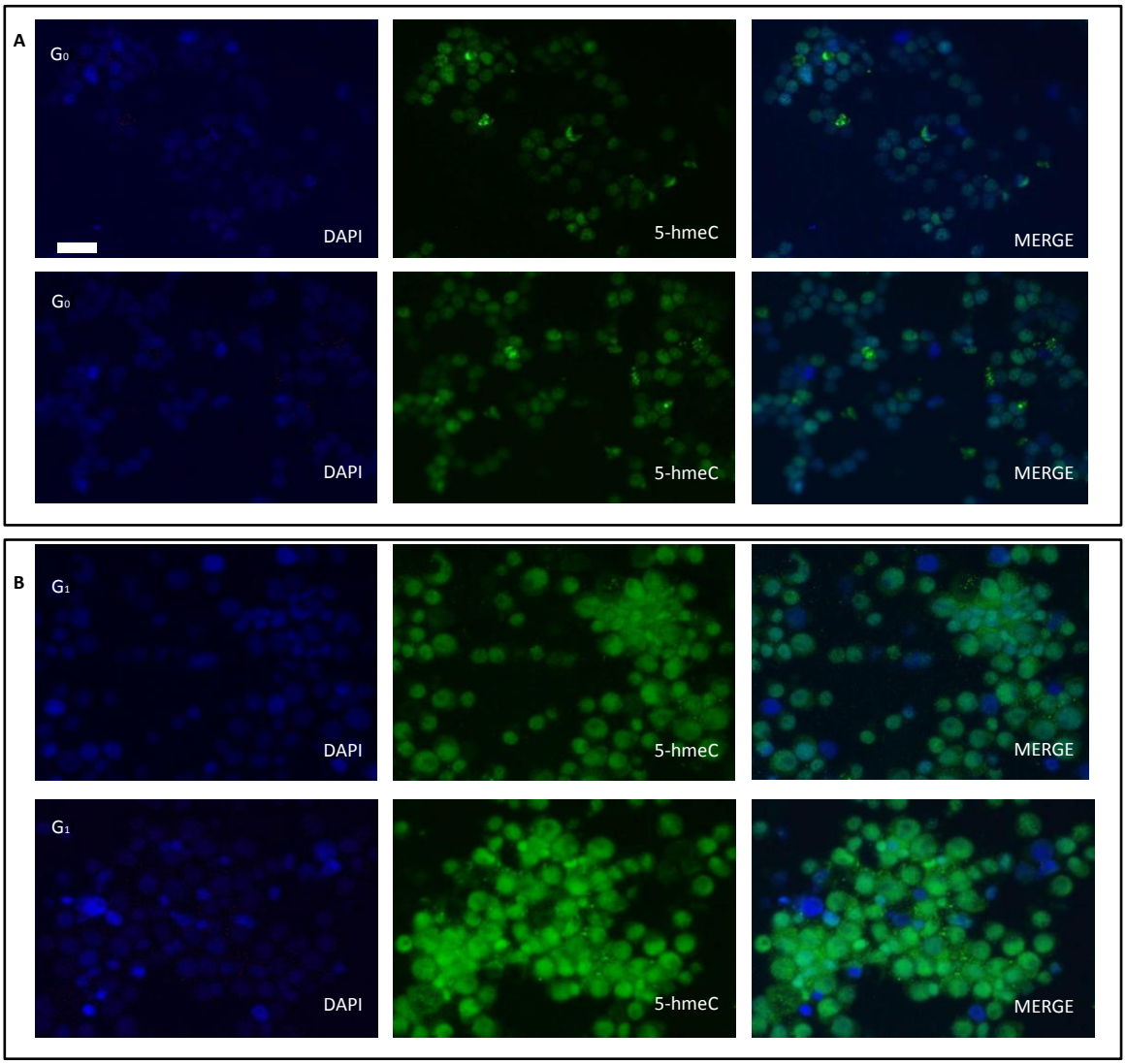


Figure 5-6 IF analysis G_0 –G_1 5-hmeC – Alexa Fluor ® 488

A. Quiescent T lymphocytes (G_0) stained with anti 5- hmeC primary antibody, followed by FITC secondary antibody. Nuclei were DAPI stained for visualization. **B** Stimulated T lymphocytes (G_1) were stained under the same conditions. *Scale bar*, 10 μm.

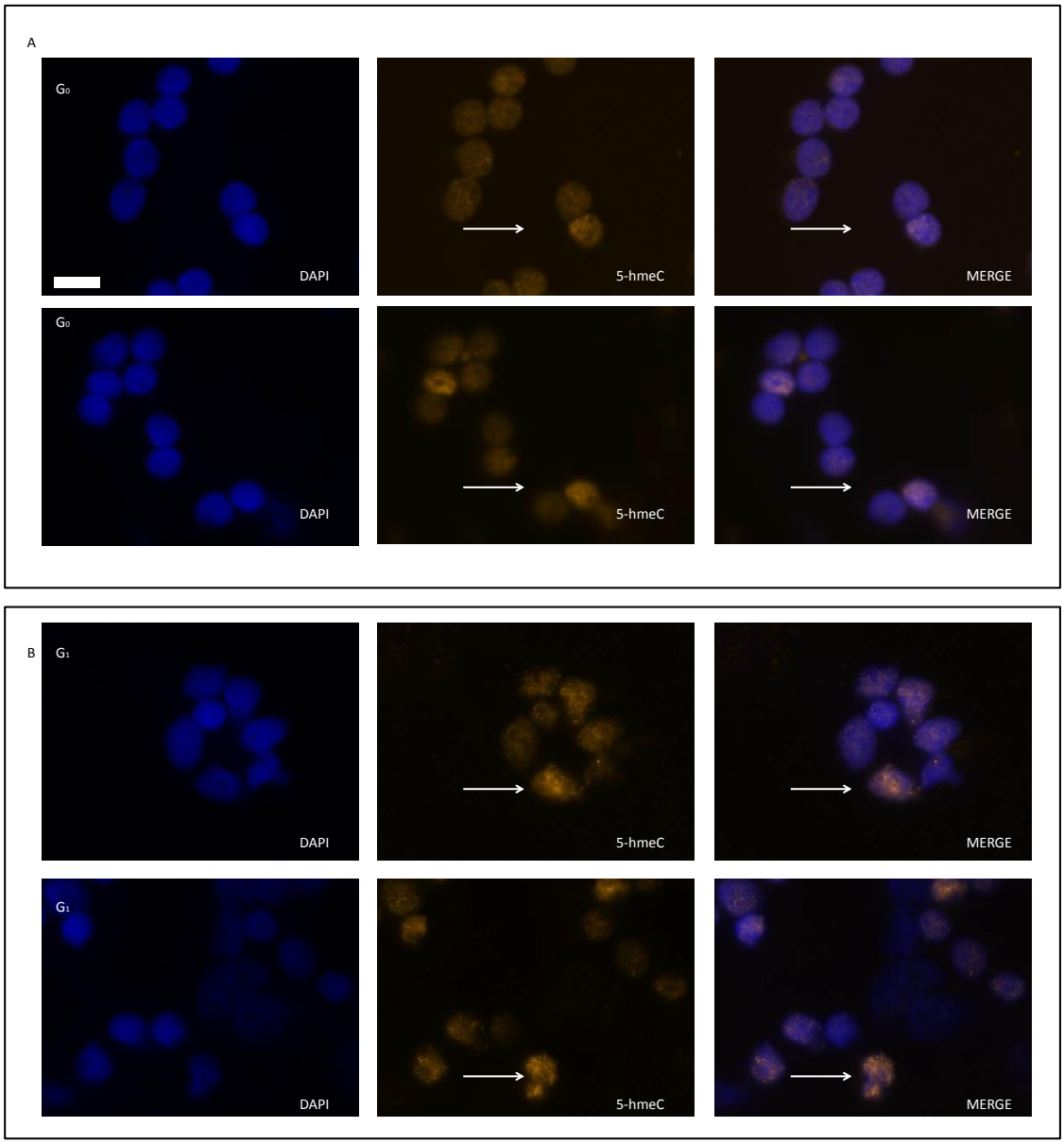


Figure 5-7 IF analysis G_0 –G_1 5-hmeC - Alexa Fluor ® 647

Quiescent T lymphocytes (G_0 stained with anti 5-hmeC primary antibody, followed by dye conjugated secondary antibody. Nuclei were DAPI stained for visualization. **B** Stimulated T lymphocytes (G_1) were stained under the same conditions. *Scale bar*, ~4 μm.

Given that the transition from G_0 to G_1 involves significant up regulation of a number of genes, either through promoter demethylation or chromatin restructuring, the observation of enhanced 5hmeC levels in G1 is perhaps unsurprising [485]. However, it could also be possible that at G_1 DNA is in much more accessible, open conformation and as such more receptive to staining compared to a more compact closed state during G_0. It has been demonstrated that access by transcriptional regulators and DNA replication mediators requires alterations in chromatin compaction to occur [486] [487]. Furthermore, recent evidence shows that 5mC is under-represented at gene promoters and CGIs whereas 5-hmeC is enriched and is associated with increased transcriptional levels[485].

To address this issue an attempt was made to carry out methylcytosine immunofluorescence to look at the DNA methylation changes during G_0-G_1 transition under the same conditions as for 5-hmeC staining. If DNA accessibility is an issue, then it would be resolved by this approach. The primary antibody used was mouse monoclonal IgG1anti-methylcytosine antibody (Eurogentec Catalogue No BI-MECY-0100). Unfortunately, no signals could be obtained even following extensive optimization procedures.

The only way to address whether these are real changes to the global levels of 5- hmeC as the cell enters the cell cycle or if the signals arise due to the experimental procedure is to use either a different meC antibody, or an anti-DNA antibody as a reference control.

However, given the observed differences in global 5-hmeC content between G_0 and G_1 cells, as real, one would also perhaps expect to see differences in TET family expression between the G_0 and G_1 status of cells.

5.3.3 TET Expression levels change during G_0-G_1 transition

Having possibly demonstrated cell cycle dependent differences in 5-hmeC signal intensity and taking into account what is currently known about the role of TET proteins in DNA hydroxymethylation I wanted to investigate further the expression profiles of TET proteins during G_0 –G_1 transition. Samples of quiescent and stimulated T lymphocytes (the same as the ones used in the IF study) were used for RNA extraction and expression profiling. Expression of TET family members was assessed using gene specific TET primers for TET1, TET2 and TET3. The endogenous reference for the PCR reaction was the 18S housekeeping gene. SYBR Green qPCR was used for relative quantification analysis of TET expression levels relative to the 18S reference. SYBR Green is a fluorescent dye which binds double–stranded DNA and is detected by measuring the increase in fluorescence throughout the cycle. RT-qPCR Assay set up, and relative quantitation using ΔΔCt method are described in Methods Chapter 2 Section 2.13.1. The outcome of the relative quantitation RT-qPCR experiment is shown in Figure 5-9A.

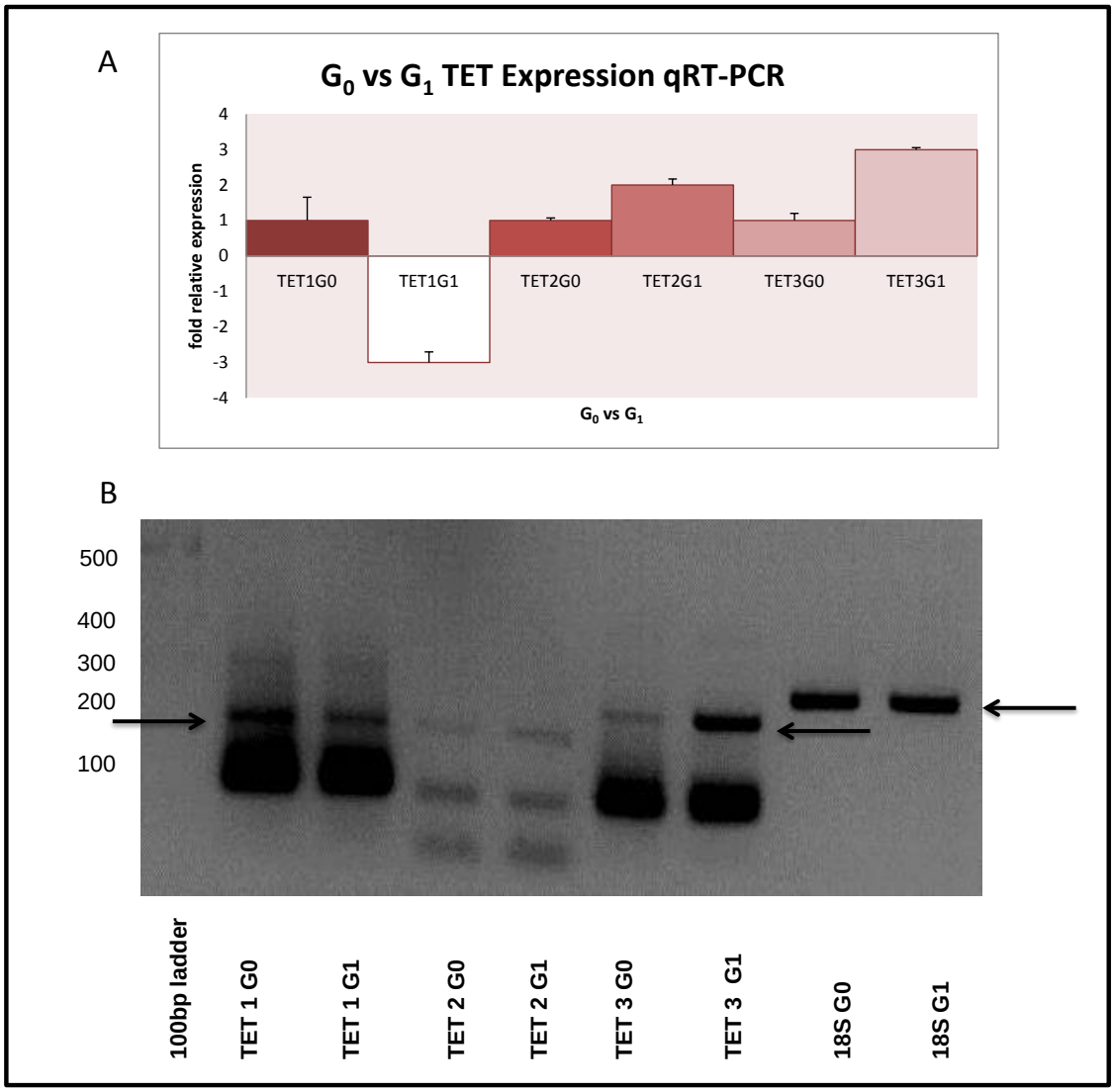


Figure 5-9 RT-qPCR TET Expression Evaluation during G_0–G_1 transition normalised to 18S

TET family member's expression profile by RT-qPCR Sybr Green relative quantitation method was used to evaluate TET1, TET2 and TET3 expression changes during the cell cycle entry. **A** RT-qPCR expression data for each of the TETs at G_0 *vs.* G_1 shows downregulation of TET1, and up regulation of TET2 and TET3 as the T lymphocytes enter the cell cycle. **B**. Agarose gel of the PCR run using HotStarTaq Polymerase shows the same result. The visual confirmation of qPCR using HotStarTaq plus DNA polymerase (Qiagen 203601) was employed (Figure 5-9B). The resulting amplicons (arrowed) were of sizes 18S - 187bp, TET1 and TET2 145bp and TET3 – 168bp.

The results of relative quantitation analysis indicate changes in TET family expression patterns during the G_0–G_1 transition. Results obtained by PCR were comparable to the RT-qPCR data. The data generated so far indicates that there is an increase in 5-hmeC signals during the cell cycle entry. The transcriptional up regulation of TET2 is in line with what is known about the role for this protein in 5-hmeC generation [488]. The significance of TET3 up regulation is less clear at this time.

TET1 downregulation also appears to correlate with what is currently known about TET1 expression patterns. Recently, studies in mouse ES cells using chromatin immunoprecipitation coupled with high-throughput DNA sequencing show that TET1 preferentially binds to CpG-rich TSSs, a significant proportion of Polycomb group (PcG) target genes and that it also associates and co localizes with the SIN3A co-repressor complex. The same study also suggests a role for TET1 in transcriptional repression of PcG targeted developmental regulators based on the observation that recruitment of PRC2 to CpG-rich gene promoters appears to be facilitated by TET1 [489]. Not much is known about the TET3 role in T cells cell cycle progression. It is particularly interesting that the highest up regulation is seen for this protein. TET3-mediated DNA hydroxylation was shown to have a role in epigenetic zygotic paternal DNA reprogramming during natural fertilization [490].

I had planned to investigate the methylation status of the *CDKN2B* promoter in response to knockdown of TET family members through siRNA action in G_0 *vs.* G_1 T cells. Due to time pressures however, this experiment could not be completed.

5.4 Conclusion

Epigenetic gene silencing associated with aberrant *CDKN2B* promoter hypermethylation patterns is one of the hallmarks of hematopoietic malignancies such as MDS and AML.

This divergent gene is normally expressed in a lineage restricted manner and at very low levels. *CDKN2B* regulates cellular senescence, hematopoietic cell differentiation and cell cycle control. Under normal physiological conditions, the *CDKN2B* promoter region is methylation free, but in disease it yields to aberrant and progressive methylation of its CGI. The mechanisms which this and other gene promoters might employ to keep their regulatory regions methylation free are unknown at this time, as are the routes by which these mechanisms might be circumvented in disease cells. In this chapter, two distinct mechanisms of protection against widespread promoter methylation in the *CDKN2B* region were investigated.

The first part of the chapter investigated a postulated role for VEZF1 as a barrier against DNA methylation activity. An interesting candidate, VEZF1, which was recently proposed to mediate protection from *de novo* DNA methylation at the chicken β-globin domain locus and the APRT CGI in ES cells [491] was an attractive model for investigation. This transcription factor was demonstrated to mediate global DNA methylation levels particularly at Line1 elements and minor satellite repeat elements, in studies involving VEZF1 -/- mouse ES cells. Furthermore, deletion of VEZF1 was accompanied by DNMT3B depletion [492]. Given that VEZF1 recognizes similar G-rich sequences as Sp1, which is a known regulator of *CDKN2B* transcription, this chapter attempted to investigate the hypothesis that VEZF1 may play a role in providing a defence against DNA methylation.

In this chapter, I showed using ChIP pull downs that VEZF1 is recruited at the 5' region of *CDKN2B* promoter. This identifies a previously unknown interaction which will hopefully advance our understanding of *CDKN2B* regulation. To investigate whether VEZF1 plays a role in protecting the *CDKN2B* promoter against 'casual' methylation repeated knockdowns of VEZF1 were performed using VEZF1 siRNA in HeLa cells. This resulted in the unexpected finding that

the $p15^{INK4B}$ specific band identified by western blot analysis disappears, and a 10kDa non-specific product co incidentally appears.

In the second part of the chapter, I considered an alternative mechanism for the maintenance of the methylation – free state at the promoter region of the transcribed genes which involves removal of methylated cytosines independent of cell division [493]. There are a number of proposed demethylation mechanisms ; base excision repair (BER) to remove 5meC usually via a DNA glycosylase, nucleotide excision repair (NER) mediated by the GADD45 family, deamination of 5meC to T via AID, APOBEC, or DNMTs, followed by BER to correct the T·G mismatch [494]. The TET family of enzymes is another group of putative demethylases. The TET1 family member was first identified as a fusion with H3K4 methyltransferase MLL in AML, but also there is now an increasing body of evidence implicating TET1 and TET2 in a haematopoietic malignancies [495].

Of particular interest was the recently proposed role for TET protein in mediating demethylation by 5meC hydrolysis to produce an intermediate hydroxymethylcytosine 5-hmeC. The ultimate aim of this research here was to evaluate the effects of TET knockdowns on *CDKN2B* promoter methylation patterns.

Cultured cell lines such as HeLa were shown to possess very low, almost undetectable levels of 5-hmeC [496], so I decided to use primary cells as a model system. T cells seemed an ideal choice, firstly because methods for cell isolation from peripheral blood, and stimulation from its native quiescent state (G_0) to enter the cell cycle following CD3/CD28 co-stimulation (G_1), were already established in our lab. Significant differences in TET family member transcription levels between the two cell cycle states were observed. Furthermore, according to preliminary immunofluorescence data, the 5-hmeC signals between the two states also differ, with an increase in 5-hmeC signal intensity at the G_1 stage of the cell cycle. This is the stage characterized by cellular growth, restructuring of surface / secreted effector function molecules and an increase in gene expression by regulatory proteins. If the 5-hmeC data are true, and not a result of an experimental ambiguity, then this would be the first representation of global demethylation being observed as the cell enters the cell cycle. However, much more work is required before we can advance, or discount this idea.

Chapter 6

General Discussion

6.1 Endogenous gene promoter region DNA Methylation targeting

Compared to the genome, the epigenome is in a forever fluid state, modelled and shaped by a variety of contributing factors, one of which is DNA methylation. Large scale DNA methylation studies in mammalian systems have shown a genome wide methylation distribution, with the conspicuous exception of CGIs. The proposed function for DNA methylation is generally accepted to be involvement in gene promoter silencing and the maintenance of the repressed state[497].

Despite the consensus view on the role of the DNA methylation in irrevocable transcriptional silencing, there is a little evidence to directly link DNA methylation to gene expression modulation. Furthermore, emerging evidence of a dynamic pattern of promoter methylation, which varies according to the cell type, and observation of the DNA methylation targeting to the transcriptional units of actively transcribing genes, highlights the limits of our understanding of this important epigenetic mechanism.

As previously mentioned epigenetic studies of the mammalian genome so far, have helped paint a picture of the DNA methylation landscape, featuring global hypermethylation, interrupted by short un-methylated CpG rich regions regions– CGIs as a most prominent feature. These CGI regions are proposed as epigenetic ON/OFF switches which exist in two alternative states; heavily methylated (repressed) and essentially unmethylated (active). Active promoter regions are shown to be methylation free in the region of 500bp each side of TSS. DNA methyltransferases are presumed to be excluded from this region by promoter bound factors [498].

Cancer and other human diseases have been linked to aberrant methylation patterns, frequently characterized by global hypomethylation, promoter hypermethylation and silencing of the key tumour suppressor genes [499]. Similar methylation patterns are observed as healthy cells age, most likely contributing to age related diseases.

Despite numerous studies investigating a co-relation between DNA methylation and gene silencing, the exact mechanisms of this relationship *in vivo* are not fully understood. One of the

longest standing questions in the field remains unanswered; whether DNA methylation initiates or is a consequence of gene silencing. By attempting to mimic DNA *in vivo* methylation events at a gene promoter region we proposed that it might be possible to perhaps dissect the temporal aspects and mechanisms of this relationship.

This project approach to the above challenge was to use zinc finger peptide and DNA methyltransferase fusions in a specific attempt to try and replicate the kind of methylation patterns observed in MDS (reviewed in Chapter 1 Section 1.4). The six finger proteins used for this project were capable of recognizing 18 bp contiguous DNA sequence. Although unknown to occur in nature, the ZFs of this size were previously shown to be effective in gene regulation.

Efforts in our lab to examine the effects of de novo methylation on gene regulation so far, have involved the use of site –biased DNA methyltransferase to deliver CpG methylation directly to the target locus via zinc finger protein technology. Using this approach it was possible to demonstrate that the targeted DNA methylation to an integrated promoter leads to gene silencing, mediated by induced repressive chromatin signature [500]. This current project was to apply a similar methodology to target a disease relevant endogenous gene promoter, in order to better understand spatial and temporal aspects of genuine *de novo* DNA methylation in *in vivo* settings.

The gene selected for this study was the *CDKN2B* tumour suppressor gene, which has a key role in cell cycle regulation, a haematopoietic cell differentiation, and cellular senescence (reviewed in Chapter 1). The gene is highly implicated and epigenetically silenced in various haematological malignancies such as MDS and AML. Aberrant methylation patterns and gene silencing of the *CDKN2B* gene have been reported, and extensively characterized for MDS. In fact, *CDKN2B* promoter methylation is currently used as a biomarker for the disease progression towards AML. *CDKN2B* suggested itself as an excellent test system for both laying down de novo methylation patterns, and assessing the functional outcome of such targeted methylation.

During this PhD study I have attempted to investigate the effects of targeting different CpG methylation densities to the promoter region of the endogenous gene, which included analysis of gene expression changes and spread of de novo methylation over time.

Below I will discuss the main findings of my PhD project, and how they relate to what is currently known about the delivery of DNA methylation to the target region, methylation spread, and its effect on gene transcriptional control.

6.1.1 Zinc Finger Protein/Methyltransferase Fusion characterization

Rationally designed zinc finger proteins fused to DNA methyltransferases were used in order to attempt delivery of low, medium and high density methylation to the *CDKN2B* key regulatory promoter region in HeLa cells, in an attempt to simulate the epigenetic events frequently observed in haematological malignancies.

The rationale for the design of the zinc finger proteins was based on extensive literature searches for key *CDKN2B* transcriptional regulators, coupled with a TF search software package.

CDKN2B specific 6ZFP proteins, fused to the HpaII FH (F35H) prokaryotic methyltransferase were used in titrative EMSA gel shift assays for comparative relative binding affinity assessments of each fusion protein. High affinity binding for two of the zinc finger fusions (PU1/HpaII and AB/HpaII) to probes harbouring the relative recognition sites and the HpaII sites was confirmed. Analysis of the binding data showed that the PU1/HpaII bound to its target site most strongly (apparent Kd 5.2nM); the AB/HpaII protein binds with intermediate strength (Kd 11nM). The Miz/HpaII fusion protein bound relatively weakly to its target site (Kd 43nM), and Myc/HpaII bound very poorly indeed (Kd 992nM). The same fusion proteins were incubated with the target oligonucleotide harbouring a scrambled zinc finger site with the HpaII binding site left intact, and the binding was observed only at the highest concentrations of proteins used. These results suggested that the zinc finger dominates binding. These results compare favourably to other studies of six finger proteins. The Kd values observed for PU1/HpaII and AB/HpaII are in a comparable range to similar six finger proteins previously characterized in our laboratory (Kd 2nM). Interestingly, the Kd of PU1/HpaII used in this study (Kd 5.2nM) is lower than zinc finger only proteins lacking methyltransferase component (Kd 7nM) previously characterized in our

lab. This observation is in line with the data, which shows that the presence of the methyltransferases subunit increases the overall binding affinity of ZF/MT fusions for substrate harbouring both recognition sites. However, as already highlighted earlier, these Kd values are guidelines only because the issues of variation in protein and DNA concentration measurements and the purity of the fusion proteins have to be taken into the account.

Given that the most of ZFs studied in these assays were designed and synthesized using the same approach the differences in their binding affinities are surprising. The modular nature of ZFs which enables to recognize the triplet DNA subsites can adversely or sometimes even positively be affected by target site overlap (TSO) interactions[501].

In vitro characterization of generated fusion proteins examined their ability to methylate the target site against a complex backdrop of HpaII sites. Plasmids harbouring a zinc finger and HpaII site target sequence were used to assess the *ex vivo* enzymatic activity of the purified proteins. A unique restriction pattern suggestive of methylation at the target site should arise following HpaII restrictions of the plasmid. In this way, it is possible to comparatively assess the enzymatic activity of various ZFP/ methyltransferases expressed in *E.Coli*.

The *ex vivo* enzymatic activity binding assays have been used previously in our lab, but inconsistencies have been noted. The attempts to assess the enzymatic activity of the purified fusion proteins were unsuccessful. Gel analysis of the purified proteins did show minor issues associated with protein degradation. It is possible that degradation or cleavage products of the purified protein, possibly resulting from the purification process, interfered with the binding assays; for example by blocking access of the full length protein.

These issues of protein degradation could possibly be addressed by changing the purification system. An alternative purification system was evaluated previously in our lab. Using the Maltose Binding Protein MBP, a greater amount of degradation was observed, suggesting no advantage to using this system (Wong.,K unpublished data). Another possibility was that the *ex vivo* assay was being influenced by the topology of the plasmid vectors used. Typical plasmid topology is a covalently closed circular (ccc) supercoiled form, although it can also appear in damaged structures, like open circular (oc) or linear forms, where one or both strands are broken, or can exist in larger concatemeric forms.

Prior to targeting the gene promoter region in its endogenous setting, which was the primary goal of this thesis, we wanted to validate these targeted methyltransferases in a *CDKN2B* relevant model system. To this end, the *CDKN2B* promoter containing primary regulatory regions and CGI was used to assess delivery of de novo CpG methylation by examining changes in reporter gene expression.

This builds on the previous published work in our lab, where the four finger (4aFH) driven targeting using the HpaII FH mutant methyltransferase, to the artificial chromosomally integrated target promoter led to de novo methylation of the target region, accompanied by the reporter system gene silencing, and the acquisition of repressive chromatin signature-H3K9me2 methylation[500].

HeLa cells were transiently co-transfected with the pBL-CAT plasmid harbouring the *CDKN2B* promoter region driving the expression of CAT reporter gene, and various CMV driven ZFP/methyltransferase fusion constructs.

CAT Activity was measured principally by CAT-ELISA. Radioactive CAT assays were also performed to confirm the results. Although relatively high transfection efficiencies were consistently achieved (80%) as confirmed by GFP fluorescence-activated cell sorter analysis, only very weak CAT signals, insufficient for further analysis were detected. The *CDKN2B* promoter on its own appears to be too weak to facilitate sufficient gene expression.

To address this issue an alternative reporter system using GFP was evaluated to more readily measure promoter activity. It was believed that it would be more sensitive than ELISA based approaches given the GFP FACS readout, but no fluorescent signals could be detected upon HeLa cells transfections confirming the previous observation of the *CDKN2B* promoter strength. Another possible explanation for the observed lack of activity could be that *CDKN2B* requires specific stimuli for its induction, or that some key regulatory sequences are missing from the promoter region used in these assays. It is noteworthy that in our hands RT-qPCR *CDKN2B* analysis showed that this is a low abundance transcript (Ct 32), whose transcription improved significantly upon phorbol myristyl acetate (PMA) treatment (Ct 28).

The *CDKN2B* gene is induced in response to TFG-β signalling via SMAD phosphorylation [502]. Furthermore, basal promoter activity is shown to increase dramatically upon activation with tetradecanoyl phorbol acetate (TPA) acting in concert with Snail, EGR-1 and AP-1 transcription factors. The TPA responsive element is located between -226 and 80bp relative to TSS, and is a binding site for SP-1 and EGR-1 TFs [503]. We have contemplated TPA treatment to induce the *CDKN2B* transcription levels but decided against it as it would have an effect on the endogenous *CDKN2B* levels, and quite likely mask the promoter/reporter signals. Several cell lines with *CDKN2B* deletions K562, Jurkat and A549 were evaluated to test whether low expression was a cell type specific issue, but their transfection efficiencies were suboptimal for any further analysis.

6.1.2 Endogenous *CDKN2B* methylation targeting

Both prokaryotic and eukaryotic methyltransferases (catalytic domain) zinc finger fusions have been successfully used to direct targeted methylation in bacteria, yeast and the mammalian mitochondrial compartment [504-506]. The effects of methylation targeting on gene expression modulation have also been assessed in transient assays involving viral vectors, plasmid based reporter genes [507] and most recently in our lab on an integrated “artificial“ target region [500]. To the best of our knowledge, no-one has yet successfully targeted de novo methylation to the more functionally complex endogenous promoter in in vivo settings.

The promoter region of an actively transcribing gene is a highly dynamic and complex region. The core promoter region serves as a docking template with recognition sequences for numerous epigenetic and transcription factors which co-ordinate transcriptional regulation. This fact, coupled with the issues of DNA topology changes resulting from various cellular processes, mechanical stresses and DNA:nucleosome interactions during the highly dynamic process of active transcription [508] made it questionable whether the specific core promoter architecture in in vivo settings could be accessed by targeted methyltransferase at all.

The benefits of successful delivery of targeted de novo methylation are manyfold. By accessing and methylating the region it would be possible for the first time ever to look at the broad

contribution this epigenetic modification makes to overall transcriptional regulation. Secondly, this approach could allow us to examine the dynamics of methylation spread and the factors associated with epigenetic inheritance. Finally, by studying the time dependant consequences of targeted methylation, important insights into the spatial and temporal relationship between the promoter gene DNA methylation and gene regulation could be gained.

An important point to reiterate at this stage is that it is still not clearly defined if DNA methylation spreading into the promoter region initiates gene silencing by recruiting repressive complexes, or if a diminishing transcription leads to the gradual increase in methylation levels to the point when stable silencing is sealed by a permanently methylated promoter. A number of studies demonstrate the existence of stably inactivated methylated promoters, where the silencing is probably mediated by MeCP-1 binding, and recruitment of repressive elements [509, 510]. However, there are studies also to support the idea of a gradually diminishing gene expression, that is not initially associated with DNA methylation but leading to a complete silencing 'sealed' by a methylated promoter. One such study involved using E6 papillomavirus transfected human mammary epithelial cells which escape senescence in correlation with gradual p16 downregulation in the absence of promoter methylation. Following a number of cell passages, complete silencing occurs associated with promoter methylation [511].

Using either single or combinatorial expression of targeted prokaryotic methyltransferases we attempted to lay down low density methylation at the target region, and investigate possible effects on the progressive methylation spread and gene transcriptional regulation.

To position low density methylation two prokaryotic methyltransferases were used; M.HpaII (F35H), which recognizes the C**C**GG sequence and M.HhaI which recognizes GCGC, both of which modify the internal (bold and underlined) cytosine residue (C5) in the sequence. M.HpaII (FH), previously used in methylation studies in our lab, is a mutated version with reduced binding affinity, thus allowing the ZF component to dominate in functional binding interactions, and effectively eliminate non-targeted background methylation [512]. For higher density methylation studies, CpG specific prokaryotic M.SssI, and the C-termini of two mammalian methyltransferases DNMT3A and DNMT3B were evaluated. I also evaluated co targeting of these enzymes with their respective co-stimulator - DNMT3L.

Accessing the actively transcribing, methylation free promoter, with its open chromatin conformation was expected to be an issue. Therefore, I was unsure whether accessing an actively transcribing region would be an issue or not. It has long been speculated that the open chromatin established by an active promoter actively blocks the spread of methylation from methylated distal region.

Here, it is demonstrated for the first time that it is possible to lay down a *de novo* methylation pattern to a defined region of an actively transcribing endogenous gene. Using M.HpaII and M.HhaI methyltransferases, specific de novo methylation was deposited at the target sites, as determined by bisulfite methylation sequencing of the library, followed by TOPO-TA cloning of individual clones for a more detailed analysis. Short-term studies (up to day 21) on the spread of *de novo* methylation from the seed target foci to the surrounding methylable CpG sites showed that there is no obvious spreading to the flanking CpG sites. Long –term studies to address the temporal and spatial methylation spread are on-going.

The analysis of the region closer to the ZF target site revealed an obvious spread of methylation towards the 3' region closer to the zinc finger target site. At Day 14 an increase in the level of methylation of around 29% was observed in this region of the promoter in the immediate vicinity of ZFs.

A second very important observation was clear evidence for epigenetic inheritance of the low density *de novo* methylation patterns. Once acquired, the methylation patterns are propagated through successive cell divisions long after the initial methylation stimuli of exogenous targeted DNA methyltransferases are lost. The fact that we demonstrated maintenance of methylation at the HpaII target site, but no subsequent methylation spread onto the local CpG site, implies de novo endogenous activities of DNMT3A or 3B.

More distal regions of the target site were defined by highly specific target site methylation, indicating that M.HpaII and M.HhaI show distinct sequence preference.

Following targeted promoter methylation, transcriptional regulation assessment of *CDKN2B* showed no change in the gene expression levels. Again, it should be pointed out that this is a

short – term study, but so far the data point towards exclusion of promoter DNA methylation as the primary causative event in gene repression.

Having established that imposed low density methylation in the *CDKN2B* promoter region was an inherited mark but with no apparent effect on gene regulation, we then asked, what would be the effect of targeting high density *de novo* methylation to that same region. To do this we employed two C-terminal mammalian methyltransferases DNMT3A and DNMT3B shown to be catalytically active [513] alone or co- targeted with DNMT3L, and a prokaryotic methyltransferase M.SssI. These enzymes preferentially methylate CpG sites and given the high abundance of these sites in the *CDKN2B* promoter target region – highly dense methylation patterns were expected. The prokaryotic CpG methyltransferase M. SssI was previously used as a Zif268 fusion, to deliver methylation to pre- determined sequence [514]. DBD fused C-terminal DNMT3A and DNMT3B have also been successfully used for targeted methylation of different reporter systems leading to gene repression.

The rationale behind co –targeting with the catalytically inactive protein DNMT3L is due to its demonstrated stimulatory effect on the catalytic activity of DNMT3A and DNMT3B [428]. Additionally, DNMT3L in heterodimer formation with DNMT3A and DNMT3B guides CpG site recognition in a periodic pattern [429] and attenuates formation of more uniform methylation patterns [430].

I was not successful in delivering high density DNA methylation using the catalytic domains DNMT3A or DNMT3B methyltransferases, with or without DNMT3L. The apparent lack of success may point to issues of enzymatic activity in our model system. It is possible that the six finger protein used in this instance binds with too strong an affinity to target site, overriding the enzymatic activity of the target methyltransferase. Another possibility is that the catalytic domain of these enzymes interact with other proteins, for example endogenous DNMT3L. Once bound, the DNMT3L coupled with DNMT3A or DNMT3B was shown to direct DNA methylation to genomic regions lacking H3K4me3[515]. Given that the *CDKN2B* promoter is transcriptionally active , it is possible that in endogenous settings targeted methyltransferases are “shuffled away” from the zinc finger target sites. This might make our decision to co-express with DNMT3L counter -productive.

Another possibility is that although we used catalytic domain methyltransferases shown to be able to methylate their target CpG substrate, additional domains of the enzyme are required for de novo DNA methylation in endogenous settings. Recently, it was shown that the ADD domain binds unmethylated H3K4 tail to preferentially methylate DNA marked by this modification. This is in line with the global analysis of DNA methylation pattern distribution [516].

It has been demonstrated that different domains of methyltransferases provide a link to mediate cross –talk between DNA methylation and chromatin signatures [517]. First shown for DNMT3L, followed by DNMT3A and DNMT3B, the PHD domain interaction with the amino-terminal tail of histone H3 is inhibited by H3K4 methylation - a marker of active genes [515, 516]. Additionally, DNMT3A via its PWWP domain interacts with the repressive modification H3K36me3, and increases the enzymatic activity for methylation of nucleosomal DNA[518]. Therefore, it appears that there has to exist an interplay between histone modifications and DNA methylase complexes to contribute to the de novo methylation process. An additional very important finding from the same study was that the N-terminus of H3 binding to the PHD domain of DNMT3A, increased its activity. DNMT3A carrying point mutations that diminished H3 tail binding but not its enzymatic activity, transfected into DNMT3A -/-, DNMT3B-/- mouse ES cells could upon differentiation bind to, but not methylate Oct4. This is the strongest indicator yet that the DNMT3A enzyme requires an interaction with H3 for its allosteric activation, and perhaps can be one possible explanation for its failure to methylate in our study. According to this scenario DNMT3A through its interactions with DNMT3L, or via PWWP, gets recruited to the chromatin template, but can be activated only through its PHD/H3K4 (unmethylated) contacts. If this is the case a much more sophisticated approach for DNMT3A and perhaps DNMT3B delivery for targeted methylation is required.

6.2 Mechanisms of *CDKN2B* promoter regulation

The work described here is the first demonstration of targeted methylation at the promoter region of the endogenous gene. Data generated so far, has not clearly suggested that promoter methylation may be an early event, that leads to *CDKN2B* gene repression frequently observed

in malignant haematopoiesis. This has left a number of open –ended questions that require further investigation.

How does the above observation fit in with the demonstrated aberrant promoter methylation patterns and associated *CDKN2B* gene silencing observed in disease? Furthermore, given that we accessed the *CDKN2B* promoter to target methylation for the purposes of this project, what is it that protects a promoter from routinely being methylated? Why does it succumb to methylation in disease? Is it possible that sustained epigentic insult in a stepwise fashion silences this gene, and that the DNA methylation is a final seal on its fate?

To attempt to begin answering some of the questions raised above we should start with taking into consideration some recent observations linking *CDKN2B* promoter DNA methylation with other components of the epigenetic silencing machinery.

CDKN2B-AS also known as ANRIL (anti-sense non-coding RNA in the INK locus), runs in the opposite direction from the INK4A-ARF-INK4B locus, and is believed to recruit PRC2 component SUZ12 leading to repression of *CDKN2B* [519]. DNA methylation fits into the model through its demonstrated physical interaction with components of the PRC1 and PRC2 complexes, such as EZH2, CBX7 and DNMT3B [520, 521]. These findings are further reinforced by a study showing loss of DNMT3B localization and DNA methylation as a consequence of PRC eviction from the INK4b-ARF-INK4a locus, upon restoration of an active SWI/SNF complex [522].

This fits in with the observation in cancer that the Polycomb target genes are more frequently shown to be aberrantly methylated [523, 524], and the evidence for the overexpression of the p15 anti sense RNA (p15AS) [525]. The same study found an inverse correlation between the expression levels of the sense and anti - sense RNA in leukaemia, and even more strikingly heterochromatin mediated silencing of p15 in the absence of DNA methylation via p15AS in in vitro system. In vivo, in mouse ES cells upon exogenous p15AS expression p15 silencing occurred through heterochromatin formation. DNA methylation was observed as a secondary effect, following ES cell differentiation.

The expression of ANRIL is regulated by another very important factor, a chromatin insulator with barrier activity, the CTCF protein which is shown to bind upstream of the INK4A gene. Global genome analysis of CTCF binding sites in HeLa and CD4+ T cells identified cell type specific CTCF binding at these barriers, and several more binding sites in the INK4 locus [526]. Furthermore, the shRNA knock downs of CTCF were shown to dramatically reduce the expression of ANRIL and INK4 locus genes [451]. Studies on the human retinoblastoma pRb gene promoter shows that binding of CTCF can be disrupted by DNA methylation, but also that this insulator protein provides methylation protection for the CGI [527].

The *CDKN2B* promoter is normally in a methylation free state, but in disease it rapidly succumbs to silencing associated with aberrant hypermethylation. So what cellular factors, and mechanisms have a role in protecting the *CDKN2B* promoter? Different scenarios are put forward to explain such mechanisms of CGI protection from methylation.

One such model proposes a role for a combinatorial binding of TFs to keep CGIs in an unmethylated state, both in cancer and in healthy cells [528]. As mentioned above, insulators such as CTCF and chromatin barrier proteins exclude methylation from regulatory elements are another proposed mechanism. In line with this observation, an interesting candidate VEZF1, was recently proposed to mediate protection from de novo DNA methylation at the chicken β-globin domain locus and the APRT CGI in ES cells[491]. Global DNA methylation studies in the VEZF1 -/- mouse ES cells show genome wide methylation loss accompanied by DNMT3B depletion. This major loss of genomic methylation occurs at specific sites such as Line1 elements, minor satellite repeat elements, some imprinted loci, and many CGIs [492].
Our study using ChIP pull downs revealed that VEZF1 is recruited at the 5' region of *CDKN2B* promoter. This was previously unknown and hopefully represents an advance towards a better understanding of the *CDKN2B* regulation. The role of VEZF1 as a protective barrier against the methylation in this region need to be further investigated.

An alternative mechanism for the maintenance of the methylation –free state at the promoter regions of transcribing genes is that a demethylation processes remove methylated cytosines independent of cell division [493].

In this project we were particularly interested in one mechanism due to the fact that TET1 family member was first identified in a fusion with H3K4 methyltransferase MLL in AML, but also there is an increasing body of evidence implicating TET1 and TET2 in a haematopoietic malignancies [495]. It has recently been proposed that TET proteins mediate demethylation involving 5meC hydrolysis to produce an intermediate hydroxymethylcytosine 5-hmeC.

Because the cultured cell lines such as HeLa express only very low levels of 5-hmeC [496], we decided to look at the primary T cells. These cells could easily be stimulated to enter the cell cycle from its native quiescent state (G_0). The TET family (TET1, TET2 and TET3) expression, and its co-relation to the 5-hmeC levels were evaluated between these two T cell states.

Our data indicates significant differences in the TET expression levels between quiescent and cycling cells. We also observed differences in the 5-hmeC signals between the two cell cycle states. Unfortunately due to time constraints we were unable to investigate further in order to confirm these preliminary immunofluorescence observations. More work is required to confirm the 5-hmeC signal differences; before we can confidently say that the observed signals are a representation of global demethylation being observed as the cell enters the cell cycle.
This is the stage characterized by cellular growth, restructuring of surface / secreted effector function molecules and an increase in gene expression by regulatory proteins and it is likely that it could be a stage also characterized by significant demethylation upon cell cycle entry.

6.3 Future Directions

Following investigations described in this thesis several lines of research should now be further pursued.

Firstly, long term targeting studies using M.HpaII and M.HhaI need to be investigated fully with a view of understanding the temporal nature of epigenetic inheritance of methylated CpG sites. When this is done further insights could be gained about the chromatin signature associated

with the targeted promoter. This should allow better interpretation of the epigenetic process associated with methylation targeting.

Next generation of targeted enzymes to deliver high density methylation will require more focus on the mutagenesis of M.SssI to allow its expression in mammalian cells. On the other hand low density methylation studies should be carried out in concert with the knockdowns of key TFs involved in the *CDKN2B* regulation. This “promoter clearance” could possibly allow the spread of methylation following de novo methylation targeting.

The second line of research which follows from Chapter 5 of this thesis, is to confirm 5-hmeC signal validity using an anti -DNA antibody. This should remove any ambiguity caused by the differences in cellular size and/or DNA accessibility between G_0 and G_1 T cells. Co-relation between the 5 hmeC and 5meC signals needs to be determined using more specific antibodies. To conclude this work TETs siRNA knockdown, its effects on cell cycle entry, *CDKN2B* expression and methylation/demethylation, and global 5-hmeC levels will need to be further assessed.

Given the role for TET proteins in malignant haematopoiesis it would be very interesting to use zinc finger TET fusions targeted to an already methylated promoter of *CDKN2B* in Raji or KG1 cell lines. As has been highlighted in the work on TET expression presented in this thesis, TET2 given its expression pattern is probably the best candidate for the pilot studies. TET targeting demethylation to an already methylated region has potential therapeutical applications which could be further investigated.

As I conclude this thesis, I realise that more questions are raised than answers provided. Whilst humbled by this thought, it is my sincere hope that some useful insights regarding the role of DNA methylation in gene regulation have been gained.

References

1. Riggs., A, Martienssen., RA, and Russo.,VEA. *Epigenetic mechanisms of gene regulation* Cold Spring Harbor Laboratory Press, 1996 p. 1-4.
2. Allis., C, *Epigenetics.* Cold Spring Harbor Laboratory Press 2007.
3. Cropley., J, Dang., TH, Martin., DI, Suter., CM, *The penetrance of an epigenetic trait in mice is progressively yet reversibly increased by selection and environment.* Proc Biol Sci. 2012: p. [Epub ahead of print].
4. Feinberg.,A, *The epigenetics of cancer etiology.* Seminars Cancer Biol, 2004. **14**: p. 427-432.
5. Riggs., AD, *X inactivation, differentiation and DNA methylation. Cytogenet.Cell Genet*, 1975. **14**: p. 9-25.
6. Pugh, JC, and Holliday.,R, *DNA modification mechanisms and gene activity during development.* Science, 1975. **186**: p. 226-232.
7. Terranova R, Agherbi.H, Boned., A, Meresse., S, Djabali., M., *Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll.* Proc.Natl.Acad.Sci., 2006. **103**: p. 6629-34.
8. Schaefer., CB, Ooi.,SKT, Bestor,,TH, Bourc'his.,D, *Epigenetic Decisions in MammalianGerm Cells.* Science, 2007. **316**: p. 398-9.
9. Wolf, SF, Lunnen.,JDJ, KD, Friedman.,T, and Migeon., BR *Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X -chromosome: Implications for X- chromosome inactivation. Proc.Natl.Acad.Sci.*, 1984. **81**: p. 2806-2810.
10. Bruniquel.,D, and Schwartz., RH, *Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. Nat. Immunol.* , 2003. **4**: p. 235-240.
11. Ehrlich., M, *Amount and distribution of 5-methylcytosine in human DNA from different types of tissues and cells. Nucleic Acids Res.* 1982. **10**: p. 2709-2721.
12. Bird., A., Taggart., M, Frommer., M., et al, *A fraction of the mouse genome that is derived from islands of non-methylated, CpG-rich DNA. Cell*, 1985. **40**: p. 91-99.
13. Riggs., AD, and Pfeifer, GP, *X-chromosome inactivation and cell memory* Trends in Genetics, 1992. **8**: p. 169-174.
14. Chae, CB, and Choi.,YC., *DNA hypomethylation and germ cell-specific expression of testis- specific H2B histone gene.* J.Biol.Chem, 1991. **266**: p. 20504-20511.
15. Ferguson-Smith.,AC Sasaki.,H, Cattanach.,BM and Surani.,MA *Parental-origin-specific epigenetic modification of the mouse H19 gene* Nature, 1993. **362**: p. 751 - 755
16. Stager.,FL, Kubicka.,P, Liu.,CG et al, *Maternal-Specific Methylation of the Imprinted Mouse Igf2r Locus Identifies the Expressed Locusas Carrying the Imprinting Signal.* 1993. **73**: p. 61-71
17. Waterland., RA, Jirtle.R., *Transposable elements: targets for early nutritional effects on epigenetic gene regulation.* Molecular and Cellular Biology, 2003 **23**: p. 5293-5300.
18. Richardson, B., *Impact of aging on DNA methylation.* Ageing Research Reviews 2003. **2**: p. 245-261.
19. Ahuja., N, and Issa.JP, *Aging, methylation and cancer.* Histol. Histopathol, 2000. **15**: p. 835-42.
20. Costello., JF, Fruhwald., MC, Smiraglia., DJ, et al, *Aberrant CpG-island methylation has non-random and tumour-type−specific patterns.* Nature Genetics 2000. **24**: p. 132 - 138.
21. Van den Wyngaert., I, Sprengel., J, Kass., SU, and Luyten., WH., *Cloning and analysis of a novel human putative DNA methyltransferase.* FEBS Lett., 1998. **426**: p. 283-9.
22. Kumar.,V, and Biswas., OK, *Dynamic state of site-specific DNA methylation concurrent to altered prolactin and growth hormone gene expression in the pituitary gland of pregnant and lactating rats.* J. Biol. Chem., 1988. **263**: p. 12645-12652.
23. Ngo., V, Gourdji., D, and Laverriere., JN , *Site-specific methylation of the rat prolactin and growth hormone promoters correlates with gene expression.* Mol. Cell. Biol. 1996. **16**: p. 3245-3254.

24. Yagi., S, Hirabayashi., K, Sato., S, et al, *DNA methylation profile of tissue-dependent and differentially methylated regions (T-DMRs) in mouse promoter regions demonstrating tissue-specific gene expression.* Genome Res., 2008. **18**: p. 1969-1978.
25. Shen., L, Kondo., Y, Guo., Y, et al, *Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters.* PLoS Genet, 2007. **3**: p. 2023-2036.
26. Kangaspeska., S, Stride., B, Metivier., R, et al.,*Transient cyclical methylation of promoter DNA.* Nature, 2008. **452**: p. 112-115.
27. Oswald., J, Engemann., S, Lane., N, et al. *Active demethylation of the paternal genome in the mouse zygote.* Curr Biol 2000. **10**: p. 475-478.
28. Mayer., W, Niveleau., A, Walter., J, et al., *Demethylation of the zygotic paternal genome.* Nature, 2000. **403**: p. 501-502.
29. Weiss., A., and Cedar., H, *The role of DNA demethylation during development.* Genes Cells, 1997. **2**: p. 481-486.
30. Bruniquel., D., and Schwartz., RH., *Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. .* Nat Immunol, 2003. **4**: p. 235-240.
31. Cantoni., GL, *The nature of the active methyl donor formed enzymatically from L-Methionine and Adenosinetriphosphate.* J.Am.Chem.Soc, 1952. **74**: p. 2942 - 2943.
32. Bestor., T, *Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain.* EMBO J, 1992. **11**: p. 2611-2617
33. Robertson., KD, *DNA methylation and chromatin - unravelling the tangled web.* Oncogene, 2002. **21**: p. 5361-5379.
34. Zimmermann., C, Guhl., E, and Graessmann., A, *Mouse DNA methyltransferase (MTase) deletion mutants that retain the catalytic domain display neither de novo nor maintenance methylation activity in vivo.* Biol Chem, 1997. **378**: p. 393-405.
35. Robertson, KD, Uzvolgyi., E, Liang., G, et al, *The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors.* Nucleic Acids Res. 1999 **27**: p. 2291-2298.
36. Bestor., T, Laudano., A, Mattaliano., R, and Ingram., V, *Cloning and Sequencing of a cDNA Encoding DNA Methyltransferase of Mouse Cells.The Carboxyl-terminal Domain of the Mammalian Enzymes is Related to Bacterial Restriction Methyltransferases.* J. Mol. Biol. 1988. **203**: p. 971-983.
37. Fatemi., M., Hermann., A, Pradhan., S, Jeltsch., A, *The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA.* J Mol Biol 2001. **309**: p. 1189-1199.
38. Pradhan., S, A.Bacolla., A, Wells., RD and Roberts., RJ, *Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation.* J Biol Chem, 1999. **274**: p. 33002-33010.
39. Hermann., A, Goyal., R, and Jeltsch., A, *The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites.* J Biol Chem. 2004. **279**: p. 48350-9.
40. Bestor, T, Laudano., A, Mattaliano., R, and Ingram., V *Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases* Journal of Molecular Biology, 1988. **203**: p. 971-983.
41. Hashimoto., H, HJ., Zhang., X, et al, *The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix.* Nature, 2008. **455**: p. 826-830.
42. Bostick., M, Kim., JK., Esteve., PO, et al, *UHRF1 plays a role in maintaining DNA methylation in mamalian cells.* Science, 2007. **317**: p. 1760-1764.
43. Kim., J, Estève., PO, Jacobsen., SE, Pradhan., S, *UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells.* Nucleic Acids Res., 2009. **37**: p. 493-505.
44. Achour., M, Fuhrmann., G, Alhosin., M, et al, *UHRF1 recruits the histone acetyltransferase Tip60 and controls its expression and activity.* Biochem Biophys Res Commun., 2009. **390**: p. 523-8.
45. Chuang., L, Ian., HI, Koh., TW, *Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1.* Science, 1997. **277**: p. 1996-2000.
46. Robertson., K, Ait-Si-Ali., S, Yokochi., T, et al, *DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters.* Nat. Genet. 2000. **25**: p. 338-42.

47. Tatematsu., K, Yamazaki., T, and Ishikawa., F, *MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase.* Genes Cells, 2000. **5**: p. 677-88.
48. Kimura., H., and Shiota., K, *Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1.* The Journal of Biological Chemistry, 2003. **278**: p. 4806-12.
49. Cardoso., MC, and Leonhardt., H, *DNA Methyltransferase Is Actively Retained in the Cytoplasm during Early Development.* J Cell Biol. 1999. **147**: p. 25-32.
50. Mertineit., C, Yoder., JA, Taketo.,T, et al, *Sex-specific exons control DNA methyltransferase in mammalian germ cells.* Development 1998. **125**: p. 889-897.
51. Takeshita., K., Suetake., I, Yamashita., E, et al, *Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1).* Proc Natl Acad Sci USA 2011. **108**: p. 9055-9059.
52. Easwaran., H, Schermelleh., L, Leonhardt., H, Cardoso., MC, *Replication-independent chromatin loading of Dnmt1 during G2 and M phases.* EMBO J, 2004. **5**: p. 1181-1186.
53. Frauer., C, and Leonhardt., H, *Twists and turns of DNA methylation.* PNAS 2011. **108**: p. 8919-8920.
54. Song., J, Rechkoblit., O, Bestor., TH, Patel., DJ ,*Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation.* Science, 2011. **331**: p. 1036-1040.
55. Sharif., J, Muto.,M, Takebayashi.,S, et al., *The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA.* Nature 2007. **450**: p. 908-912.
56. Li, E, Beard., and Jaenisch., R, *Role for DNA methylation in genomic imprinting* Nature, 1993. **366**: p. 362 - 365
57. Li, E, Bestor., TH, and Janeisch., R, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* Cell, 1992. **69**: p. 915-26.
58. Walsh., CP, Chaillet., JR, and Bestor., TH, *Transcription of IAP endogenous retroviruses is constrained by cytosine methylation.* Nat. Genetics, 1998 **20**: p. 116-7.
59. Robertson., KD, *DNA methylation, methyltransferases, and cancer.* Oncogene, 2001. **20** p. 3139-3155.
60. Pradhan.,S, and Kim., GD, *The retinoblastoma gene product interacts with maintenance human DNA (cytosine-5) methyltransferase and modulates its activity.* EMBO J. 2002. **21**: p. 779-788.
61. Chuang., LSH, Ian., HI, Koh., TW, et al, *Human DNA-(Cytosine-5) Methyltransferase-PCNA Complex as a Target for p21WAF1* Science, 1997 **277**: p. 1996 - 2000.
62. Rountree., MR, Bachman., KE, and Baylin.,SB, *DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci.* Nature Genetics, 2000. **25**: p. 269 - 277.
63. Hermann., A, Schmitt., S, Jeltsch., A, *The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity.* J. Biol. Chem. 2003. **278**: p. 31717-21.
64. Liu., K, Wang., YF, Cantemir., C, and Muller., MT, *Endogenous Assays of DNA Methyltransferases: Evidence for Differential Activities of DNMT1, DNMT2,and DNMT3 in Mammalian Cells In Vivo.* Mol. Cell. Biol. 2003. **23**: p. 2709-271.
65. Goll., MG, Kirpekar., F, Maggert., KA, et al, *Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2.* Science, 2006. **311**: p. 395-8.
66. Jurkowski., T, Meusburger., M, Phalke., S, et al., *Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism.* RNA, 2008 p. 1663-70. .
67. Aapolaa., U, Lyleb., R, Krohna.,K, et al, *Isolation and initial characterization of the mouse Dnmt3l gene.* Cytogenet Cell Genet 2001. **92**: p. 122-126
68. Okano., M, Xie., S, and Li., E, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.* Nat. Genet., 1998. **19**: p. 219-220.
69. Okano., M., Bell., DW, Haber., DA, and Li., E,, *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.* Cell, 1999. **99**: p. 247-257.
70. Chen., T, Ueda., Y, Dodge., EJ, Wang., Z, and Li., E,, *Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b.* Mol. Cell. Biol., 2003. **23**: p. 5594-5605.
71. Liang.,G, Chan., MF, Tomigahara., Y, *Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements.* Mol.Cell. Biol. 2002. **22**: p. 480-491.

72. Jeong., S, Liang., G, Sharma., et al, *Selective Anchoring of DNA Methyltransferases 3A and 3 to Nucleosomes Containing Methylated DNA.* Molecular and Cellular Biology, 2009. **29**: p. 5366-5376.
73. Chen., T, Tsujimoto., N, and Li.,E, *The PWWP Domain of Dnmt3a and Dnmt3b Is Required for Directing DNA Methylation to the Major Satellite Repeats at Pericentric Heterochromatin.* Mol. Cell. Biol. 2004 **24**: p. 9048-9058.
74. Gowher., H, and Jeltsch., A, *Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases.* J. Biol. Chem., 2002. **277**: p. 20409-14.
75. Bachman., K., Rountree., MR, Baylin., SB, *Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin.* J Biol Chem., 2001. **276**: p. 32282-7.
76. Fuks., F, Burgers., WA, Godin., N, et al., *Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription.* EMBO J., 2001. **20**: p. 2536-44.
77. Bourc'his D, and Bestor,,TH, *Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L* Nature, 2004. **431**: p. 96-99.
78. Suetake., I, Shinozaki., F, Miyagawa., J, et al, *DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction.*J Biol Chem. 2004 **279**: p. 27816-23.
79. Ooi., S, Qiu., C, Bernstein., E, Li., K, *DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.* Nature, 2007. **448**: p. 714-717.
80. Xie., S, Wang., Z, Okano, M, et al, *Cloning, expression and chromosome locations of the human DNMT3 gene family.* Gene, 1999. **236**: p. 87-95
81. Jin., F, Dowdy., SC, Xiong., Y, et al., *Up-regulation of DNA methyltransferase 3B expression in endometrial cancers.* Gynecol Oncol., 2005. **96**: p. 531-8.
82. Simão Tde., A., Simões., GL, Ribeiro., FS, et al. *Lower expression of p14ARF and p16INK4a correlates with higher DNMT3B expression in human oesophageal squamous cell carcinomas.* Hum Exp Toxicol., 2006 **25**: p. 515-22.
83. Camoriano., M., Kinney., SR, Moser., MT, et al., *Phenotype-specific CpG island methylation events in a murine model of prostate cancer.* Cancer Res. 2008. **68**: p. 4173-82.
84. Roll., J., Rivenbark., AG, Jones., WD, Coleman., WB., *DNMT3b overexpression contributes to a hypermethylator phenotype in human breast cancer cell lines.* Mol Cancer. 2008. **25**: p. 7-15.
85. Okano., M, Bell., DW, Haber., DA, and Li., E, *DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development* Cell, 1999. **99**: p. 247-257
86. Surani., M, *Imprinting and the initiation of gene silencing in the germ line.* Cell, 1998. **93**: p. 309-12.
87. Weisenberger., D, Velicescu., M, Cheng., JC, et al.,*Role of the DNA methyltransferase variant DNMT3b3 in DNA methylation.* Mol Cancer Res. , 2004. **2**: p. 62-72.
88. Chen., T, Ueda., Y, Xie., S, Li., E., *A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation.* J Biol Chem. 2002. **277**: p. 38746-54.
89. Wang., J, Bhutani., M, Pathak., AK, et al., *Delta DNMT3B variants regulate DNA methylation in a promoter-specific manner.* Cancer Res., 2007. **67**: p. 10647-52.
90. Wilson., GG, *Restriction and modification systems.* Annu. Rev. Genet, 1991. **25**: p. 585-27.
91. Wilson., GG., *Organization of restriction-modification systems.* Nucleic Acids Research, 1991. **19**: p. 2539-2564.
92. Donahue., JP, and Peek., JRM, *Helicobacter pylori: Physiology and Genetics Restriction and Modification Systems.* ASM Press, 2001.
93. Lobner-Olesen., A, Skovgaard., O, and Marinus., MG, *Dam methylation: coordinating cellular processes.* Curr Opin. Microbiol. , 2005 **8**: p. 154-60.
94. Jeltsch.,A, Jurkowska., RZ, Jurkowski.,K. TP, et al, *Application of DNA methyltransferases in targeted DNA methylation.* Appl Microbiol Biotechnol, 2007 **75**: p. 1233-1240.
95. Wu., JC, and Santi., DV, *Kinetic and Catalytic Mechanisom of HhaI Methyltransferase.* The Journal of Biological Chemistry, 1987. **262**: p. 4776-4786.

96. O'Gara., M, Horton., JR, Roberts., RJ and Cheng., X, *Structures of HhaI methyltransferase complexed with substrates containing mismatches at the target base.* Nature Structural Biology, 1998 **5**: p. 872 - 877.
97. Klimasauskas., S, Kumar., S, Roberts., RJ, Cheng., X, *HhaI methyltransferase flips its target base out of the DNA helix.* Cell, 1994. **76**: p. 357-69.
98. Jiang., YL, and Stivers.,JT, *Base-Flipping Mutations of Uracil DNA Glycosylase: Substrate Rescue Using a Pyrene Nucleotide Wedge.* Biochemistry, 2002. **41**: p. 11248-11254.
99. Wu., S, and .Zhang., Y, *Active DNA demethylation: many roads lead to Rome.* Nature Reviews Molecular Cell Biology, 2010. **11**: p. 607-620.
100. Kress., C, Thomassin., H, Grange., T, *Local DNA demethylation in vertebrates: how could it be performed and targeted?* FEBS Letters, 2001. **494**: p. 135-140.
101. Métivier., R, Gallais., R, Tiffoche., C, et al. *Cyclical DNA methylation of a transcriptionally active promoter.* Nature, 2008. **452**: p. 45-50.
102. Gjerset., R., and Martin., Jr, DW, *Presence of a DNA demethylating activity in the nucleus of murine erythroleukemic cells.* The Journal of Biological Chemistry, 1982 **257**: p. 8581-8583.
103. Weiss., A, Keshet., I, Razin., A, and Cedar., H, *DNA demethylation in vitro: involvement of RNA.* Cell 1996. **86**: p. 709-718.
104. Swisher., J, Rand., E, Cedar., H, and Marie Pyle., A, *Analysis of putative RNase sensitivity and protease insensitivity of demethylation activity in extracts from rat myoblasts.* Nucleic Acids Res, 1998. **26**: p. 5573-5580.
105. Jost., J, *Nuclear extracts of chicken embryos promote an active demethylation of DNA by excision repair of 5-methyldeoxycytidine.* Proc. Natl. Acad. Sci. USA, 1993. **90**: p. 4684-4688.
106. Cortázar., D, Kunz., C, Saito., P, et al, *The enigmatic thymine DNA glycosylase.* DNA Repair (Amst.), 2007. **6**: p. 489-504.
107. Ng., H, Zhang., Y, Hendrich., B, et al., *MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex.* Nat. Genet., 1999. **23**: p. 58-61.
108. Bhattacharya., S., Ramchandani., S, Cervoni., N, and Szyf., M, *A mammalian protein with specific demethylase activity for mCpG DNA.* Nature, 1999. **397**: p. 579-583.
109. Barreto., G., Schafer., A, Marhold., J, et al. *Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation.* Nature, 2007. **445**: p. 671-675.
110. Jin., S, Guo., C, and Pfeifer., GP, *GADD45A does not promote DNA demethylation.* PLoS Genet., 2008. **4**: p. e1000013 10.1016.
111. Ito., S, Shen., L, Dai., Q, et al. *Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine.* Science, 2011.
112. Ito., S, D'Alessio., AC, Taranova., OV, et al. *Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification.* Nature, 2010. **466**: p. 1129-33.
113. Ficz., G, Branco., MR, Seisenberger, S, et al. *Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation.* Nature, 2011. **473**: p. 398-402.
114. Bhutani., N, Burns.,DM, and Blau.,HM, *DNA Demethylation Dynamics.* Cell 2011. **146**: p. 866-872.
115. Rai., K., Huggins., IJ, James., SR, et al. *DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45.* Cell, 2008. **135**: p. 1201-1212.
116. Muramatsu., M., Kinoshita., K, Fagarasan., S, et al,, *Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme.* Cell, 2000. **102**: p. 553-563.
117. Popp., C., Dean., W, Feng., S, et al, *Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency.* Nature 2010. **463**: p. 1101-1105.
118. Bhutani., N, Brady., JJ, Damian, M, *Reprogramming towards pluripotency requires AID-dependent DNA demethylation.* Nature, 2010. **463**: p. 1042-1047.
119. Cortellino., S, Xu., J, Sannai., M, et al, *Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair.* Cell 2011 **146**.
120. Guo., J, Su., Y, Zhong., C, et al, *Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain.* Cell, 2011. **145**: p. 423-434.
121. Esteller M, Corn., PG, Baylin SB, and Herman JG, *A Gene Hypermethylation Profile of Human Cancer.* Cancer Research, 2001. **61**: p. 3225-3229.

122. Frank, D, Keshet., I, Shani, M, et al *Demethylation of CpG islands in embryonic cells* Nature, 1991. **351**: p. 239-241.
123. Bird, AP, *CpG-rich islands and the function of DNA methylation.* Nature, 1986. **321**: p. 209 - 213
124. Cramer, P, Bushnell, DA, Kornberg, RD. *Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution.* Science, 2001. **292**: p. 1863-1876.
125. Conaway, R., and Conaway, J. *General transcription factors for RNA polymerase II.* Prog. Nuc. Acid Res. Mol.Biol., 1997. **56**: p. 327-346.
126. Kim, Y, Björklund, S., Li, Y, Sayre, MH., and Kornberg, RD. *A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II.* Cell, 1994. **77**: p. 599-608
127. Taylor, I, Workman,JL., Schuetz. ,TJ, Kingston,RE., *Facilitated binding of GAL4 and heat shock factor to nucleosomal templates: differential function of DNA-binding domains.* Genes Dev., 1991 **5**: p. 1285-98.
128. Guzmán, E., Lis, JT., *Transcription Factor TFIIH Is Required for Promoter Melting In Vivo.* Mol Cell Biol. 1999. **19**: p. 5652-5658.
129. Fish, R., Kane, CM, *Promoting elongation with transcript cleavage stimulatory factors.* Biochimica et Biophysica Acta 2002. **1577**: p. 287 - 30.
130. Yamamoto, S., Watanabe,Y., van der Spek, PJ.,et al. *Studies of nematode TFIIE function reveal a link between Ser-5 phosphorylation of RNA polymerase II and the transition from transcription initiation to elongation.* Mol Cell Biol., 2001 **21**: p. 1-15.
131. Holler., M, Westin., G, Jiricny., J, and Schaffner., W, *Spl transcription factor binds DNA and activates transcription even when the binding site is CpG methylated.* Genes and Development, 1988. **2**: p. 1127-1135.
132. Bird AP, Wolffe., A, *Methylation-Induced Repression— Belts, Braces, and Chromatin.* Cell, 1999. **99**: p. 451-454.
133. Hendrich., B, Haedeland., U, Ng., HH, et al,. *The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites.* Nature,1999. **401**: p. 301-304.
134. Fatemi., M, Wade., P, *MBD family proteins: reading the epigenetic code.* Journal of Cell Science, 2006. **119**: p. 3033-3037.
135. Strahl., BD, and Allis CD, *The language of covalent histone modifications.* Nature, 2000. **403**: p. 41-5.
136. Nan., X, Ng.,H, Johnson., CA, et al. *Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex.* Nature,1998. **393**: p. 386-389.
137. Harikrishnan., KN, Chow., MZ, Baker EK, et al., *Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing.* 2005. **37**: p. 254 - 264.
138. Wolffe., AP, Matzke., M., *Epigenetics:Regulation through repression.* Science, 1999. **286**: p. 481-486.
139. Luger., K, Mader., AW, Richmond., RK, et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**: p. 251-60.
140. Thomas., JO, *Histone H1: location and role.* Current Opinion in Cell Biology, 1999. **11**: p. 312-317.
141. Luger., K, *Structure and dynamic behavior of nucleosomes.* Curr. Opin. Genet. Dev, 2003. **13**: p. 127-35.
142. Grant., PA, *A tale of histone modifications.* Genome Biol, 2001. **2**: p. reviews 0003.1–0003.6.
143. Jenuwein., T, and Allis., C, *Translating the histone code.* Science 2001. **293**: p. 1074-80.
144. Bannister., A, and Kouzarides., T, *Regulation of chromatin by histone modifications.* Cell Research, 2011. **21**: p. 381-395.
145. Becker, P, and Horz, W, *ATP-dependent nucleosome remodeling* Annu. Rev. Biochem. 2002. **71**: p. 247-273.
146. Meersseman, G, Pennings, S, and Bradbury, EM, *Mobile nucleosomes - a general behavior. .* EMBO J. 1992. **11**: p. 2951-2959.
147. Ito, T, Bulger, M, Pazin, MJ, et al. *ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor.* Cell 997. **90**: p. 145-155.
148. Tsukiyama., T, Daniel.,C, Tamkun., J, et al. *ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor.* Cell, 1995. **83**: p. 1021-1026.

149. Hamiche., A, Sandaltzopolous., R., Gdula., DA, and Wu., C. *ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF.* Cell, 1999. **97**: p. 833-842.
150. Narlikar., G, Fan, HY, and Kingston, RE, *Cooperation between complexes that regulate chromatin structure and transcription.* Cell, 2002. **108**: p. 475-487.
151. Saha., A., Wittmeyer, J., and Cairns, BR, *Chromatin remodelling: the industrial revolution of DNA around histones. .* Nature Rev.Mol. Cell Biol., 2006. **7**: p. 437-447.
152. Bazett-Jones., DP, Cote., J., Landel., CC,et al, *The SWI/SNF Complex Creates Loop Domains in DNA and Polynucleosome Arrays and Can Disrupt DNA-Histone Contacts within These Domains.* Moleular and Cell Biology, 1999. **19**: p. 1470-1478.
153. Bannister., AJ, Zegerman., P, Partridge., JF, et al., *Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain.* Nature, 2001. **410**: p. 120-124.
154. Lomberk., G., Bensi., D, Fernandez-Zapico., ME, Urrutia., R., *Evidence for the existence of an HP1-mediated subcode within the histone code.* Nat Cell Biol., 2006. **8**: p. 407-15.
155. Allfrey., VG., Faulkner., R, and Mirsky., AE, *Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis.* Proc. Natl. Acad. Sci. USA, 1964. **51**: p. 786-794.
156. Brownell., J, Zhou., J., Ranalli., T, et al., *Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation.* Cell, 1996. **84**: p. 843-851.
157. Taunton., J, Hassig., C, Schreiber., SL, *A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p.* Science, 1996. **272**: p. 408-411.
158. Sewack., GF, Ellis., TW, and Hansen., U, *Binding of TATA binding protein to a naturally positioned nucleosome is facilitated by histone acetylation.* Mol. Cell. Biol 2001. **21**: p. 1404-1415.
159. Clayton., A, Hazzalin., C, Mahadevan., L, *Enhanced histone acetylation and transcription: a dynamic perspective.* Mol Cell 2006. **23**: p. 289-9.
160. Tse., C, Sera., T, Wolffe., AP, and Hansen., JC, *Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. .* Mol. Cell. Biol.1998. **18**: p. 4629-4638.
161. Zhang., K, Faiola., F, Martinez., E., *Six lysine residues on c-Myc are direct substrates for acetylation by p300.* Biochem Biophys Res Commun. , 2005. **336**: p. 274-80.
162. Iyer., N, Xian., J, Chin., SF, et al., *p300 is required for orderly G1/S transition in human cancer cells.* Oncogene. 2007. **26**: p. 21-9.
163. Yang., X, and Seto., E, *HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention.* Oncogene. 2007. **26**: p. 5310-5318.
164. Grunstein., M, *Yeast heterochromatin: regulation of its assembly and inheritance by histones.* Cell, 1998. **93**: p. 325-328.
165. Marks., P., Xu., WS, *Histone Deacetylase Inhibitors: Potential in Cancer Therapy.* J. Cell. Biochem. 2009. **107**: p. 600-8.
166. Rea, S, Eisenhaber., F., O'Carroll, D., et al., *Regulation of chromatin structure by site-specific histone H3 methyltransferases.* Nature 2000. **406**: p. 593-599.
167. Bedford., M, and Clarke., SG, *Protein arginine methylation in mammals: who, what, and why.* Mol Cell 2009. **33**: p. 1-13.
168. Lan., F, and Shi., Y, *Epigenetic regulation: methylation of histone and non-histone proteins.* Sci China C Life Sci, 2009. **52**: p. 311-22.
169. Jenuwein., T, Allis.,CD, *Translating the histone code* Science, 2001. **293**: p. 1074-80.
170. Rea., S, Eisenhaber., F, O'Carroll., D, et al, *Regulation of chromatin structure by site-specific histone H3 methyltransferases.* Nature 2000. **406**: p. 593-599.
171. Dillon., S, Zhang., X, Trievel., RC, and Cheng., X, *The SET-domain protein superfamily: protein lysine methyltransferases.* Genome Biology, 2005. **6**.
172. Bannister., A, Zegerman., P, Partridge., JF, et al., *Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain.* Nature, 2001. **410**: p. 120-124.
173. Margueron., R., Justin., N, Ohno., K, et al, *Role of the polycomb protein EED in the propagation of repressive histone marks.* Nature. 2009. **461**: p. 762-7.
174. Bernstein., B, Mikkelsen., TS, Xie., X, et al, *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell. 2006 **125**: p. 315-26.
175. Asp., P, Bluma., R, Vethanthama., V, et al, *Genome-wide remodeling of the epigenetic landscape during myogenic differentiation.* PNAS, 2011.
176. Bannister., A, Schneider., R, and Kouzarides., T, *Histone Methylation: Dynamic or Static.* Cell, 2002. **109**: p. 801-806.

177. Shi., Y, Lan., F, Matson., C, et al., *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.* Cell, 2004. **119**: p. 941-953.
178. Perillo., B, Ombra., MN, Bertoni., A, et al, *DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression.* Science. 2008. **319**: p. 202-6.
179. Klose., R., and Zhang., Y, *Regulation of histone methylation by demethylimination and demethylation.* Nat Rev Mol Cell Biol, 2007. **8**: p. 307-318.
180. Whetstine., J, Nottke., A, Lan., F, et al., *Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases.* Cell, 2006. **125**: p. 467-481.
181. Mosammaparast., N, and Shi., Y, *Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases.* Annu Rev Biochem, 2010. **79**: p. 155-179.
182. Klose., R., Kallin., EM, and Zhang., Y, *JmjC-domain-containing proteins and histone demethylation.* Nature Reviews Genetics 2006. **7**: p. 715-727.
183. Oki., M, Aihara., H, Ito., T. *Role of histone phosphorylation in chromatin dynamics and its implications in diseases.* Subcell Biochem Biophys Res Commun., 2007. **41**: p. 319-336.
184. Nowak SJ, and Corces., V, *Phosphorylation of histone H3 correlates with transcriptionally active loci.* Genes Dev, 2000. **14**: p. 3003–3013.
185. Dou., Y, Gorovsky., M, *Phosphorylation of linker histone H1 regulates gene expression in vivo by creating a charge patch.* Mol Cell, 2000. **6**: p. 225-31.
186. Cheung., P, Tanner., K, Cheung WL, et al. *Synergistic Coupling of Histone H3 Phosphorylation and Acetylation in Response to Epidermal Growth Factor Stimulation* Molecular Cell, 2000. **5**: p. 905-915.
187. Van Hooser, A, Goodrich, DW., Allis, CD., Brinkley, BR., Mancini, MA., *Histone H3 phosphorylation is required for the initiation, but not maintenance, of mammalian chromosome condensation.* J. Cell. Sci.1998. **111**: p. 3497-3506.
188. Lo, W, Trievel, RC., Rojas, JR., et al., *Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14.* Mol. Cell. Biol. 2000. **5**: p. 917-926.
189. Choo., KHA, *Centromerization* Trends in Cell Biology, 2000. **10**: p. 182-188
190. Heo K, Kim., H, Choi., SH ,et al. *FACT-Mediated Exchange of Histone Variant H2AX Regulated by Phosphorylation of H2AX and ADP-Ribosylation of Spt16.* Molecular Cell, 2008. **30**: p. 86-97.
191. Henikoff., S, *Nucleosome destabilization in the epigenetic regulation of gene expression.* Nature Rev. Genet., 2008. **9** p. 15-26.
192. Zhang, H., Roberts, DN., and Cairns, BR. *Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss.* Cell 2005. **123**: p. 219-231.
193. Santenard., A, Ziegler-Birling., C, Koch., M, et al, *Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3.* Nature Cell Biology 2010. **12**: p. 853-862.
194. Hurd., P, Bannister., AJ, Halls., K, et al., *Phosphorylation of histone H3 Thr-45 is linked to apoptosis.* J Biol Chem., 2009. **284**: p. 16575-83.
195. Fuks., F, Burgers., WA, Godin.,N, *Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription.* The EMBO Journal 2001. **20** p. 2536-2544.
196. Fuks., F, Burgers., WA, Brehm., A, et al, *DNA methyltransferase Dnmt1 associates with histone deacetylase activity.* Nature Genet., 2000. **24**: p. 88-91.
197. Jones, PL, Veenstra., GJ, Wade., PA, et al. *Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription.* Nature Genet. 1998. **19**: p. 187-191.
198. Viré., E., Brenner., C, Deplus., R, et al, *The Polycomb group protein EZH2 directly controls DNA methylation.* Nature, 2006. **439**: p. 871-4.
199. Steensma., DP, and Tefferi., A, *The myelodysplastic syndrome(s): a perspective and review highlighting current controversies* Leukemia Research 2003. **27**: p. 95-120.
200. Pedersen-Bjergaard., J, Christiansen., DH, Andersen., MK, Skovby., F, *Causality of myelodysplasia and acute myeloid leukemia and their genetic abnormalities.* Leukemia, 2002. **16**: p. 2177-2184.
201. Jones., P, and Baylin., SB, *The epigenomics of cancer.* Cell 2007. **128**: p. 683-692.
202. Feinberg., A, and Vogelstein., B, *Hypomethylation distinguishes genes of some human cancers from their normal counterparts.* Nature,1983. **301**: p. 89-92.

203. Robertson., K, Uzvolgyi., E, Liang., G, et al, *The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors.* Nucleic Acids Res, 1999. **27**: p. 2291-2298.
204. Baylin., S, Herman., JG, Graff., JR, et al, *Alterations in DNA methylation -A fundamental aspect of neoplasia.* Adv Cancer Res, 1998. **72**: p. 141-196.
205. Sakai., T, Toguchida., J, Ohtani., N, et al, *Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. .* Am J Hum Genet, 1991. **48**: p. 880-888.
206. Myohanen., S, Baylin., SB, Herman., JG, *Hypermethylation can selectively silence individual p16ink4A alleles in neoplasia.* Cancer Res., 1998. **58**: p. 591-593.
207. Herman., J, Latif., F, Weng., Y, et al, *Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma.* Proc Natl Acad Sci U S A, 1994. **91**: p. 9700-9704.
208. Kane., M, Loda., M, Gaida., GM, et al, *Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repairdefective human tumor cell lines.* Cancer Res., 1997. **57**: p. 808-811.
209. Corn., PG, SD., Ruckdeschel., ES et al, *E-cadherin expression is silenced by 5' CpG island methylation in acute leukemia.* Clinical Cancer Research 2000. **6**: p. 4243-4248.
210. Dhodapkara., M, Grill. J, and Lust JA, *Abnormal regional hypermethylation of the calcitonin gene in myelodysplastic syndromes* Leukema Research 1995. **19**: p. 719-726.
211. Brakensiek K , LF., Schlegelberger., B, et al, *Hypermethylation of the suppressor of cytokine signalling-1 (SOCS-1) in myelodysplastic syndrome.* British Journal of Haematology, 2005. **130**: p. 209-217.
212. Quesnel B, Guillerm., G., Vereecque R, et al, *Methylation of the p15INK4b Gene in Myelodysplastic Syndromes Is Frequent and Acquired During Disease Progression.* 1998 **91**: p. 2985-2990.
213. Mizuno S, Chijiwa., T., Okamura T, et al, *Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia* Blood, 2001 **97**: p. 1172-1179
214. Lubbert., M, *Gene silencing of the p15/INK4B cell-cycle inhibitor by hypermethylation: an early or later epigenetic alteration in myelodysplastic syndromes?* Leukemia 2003. **17**: p. 1762-1764.
215. Dao., MA, Taylor., N, and Nolta JA, *Reduction in levels of the cyclin-dependent kinase inhibitor p27(kip-1) coupled with transforming growth factor beta neutralization induces cell-cycle entry and increases retroviral transduction of primitive human hematopoietic cells.* 1998. **95**: p. 13006–13011.
216. Teofili., L, Morosetti., R., Martinib M, et al, *Expression of cyclin-dependent kinase inhibitor p15INK4B during normal and leukemic myeloid differentiation* Experimental Hematology, 2000. **28**: p. 519-526.
217. Raj., K, Mufti., G, *Azacytidine (Vidaza®) in the treatment of myelodysplastic syndromes.* Therapeutics and Clinical Risk Management, 2006. **2**: p. 377-388.
218. Drexler., HG, *Review of alterations of the cyclin-dependent kinase inhibitor INK4 family genes p15, p16, p18 and p19 in human leukemia-lymphoma cells.* Leukemia, 1998. **12**: p. 845-59.
219. Raj., K, John., A., Ho., A, et al. *CDKN2B methylation status and isolated chromosome 7 abnormalities predict responses to treatment with 5-azacytidine.* Leukemia, 2007. **21**: p. 1937–1944.
220. Raj., K, and Mufti., G, *Azacytidine (Vidaza®) in the treatment of myelodysplastic syndromes.* Ther Clin Risk Manag, 2006. **2**: p. 377–388.
221. Lehmann., U, Brakensiek., K., Kreipe., H, *Role of epigenetic changes in hematological malignancies.* Ann Hematol 2004. **83**: p. 137-152.
222. Tessema., M, Langer., F, Dingemann J, et al, *Aberrant methylation and impaired expression of the p15INK4b cell cycle regulatory gene in chronic myelomonocytic leukemia (CMML).* Leukemia, 2003 **17**: p. 910-918.
223. Tien HW, TJ, Tsay., W, et al, *Methylation of the p15INK4B gene in myelodysplastic syndrome: it can be detected early at diagnosis or during disease progression and is highly associated with leukaemic transformation.* British Journal of Haematology, 2001. **112**: p. 148-154.
224. Christiansen DH, AM., and Pedersen-Bjergaard J., *Methylation of p15INK4B is common, is associated with deletion of genes on chromosome arm 7q and redicts a poor prognosis in therapy related myelodyspasia and acute myeloid leukemia. .* Leukemia, 2003. **17**: p. 1813-1819.

225. Lehrmann., H., Pritchard., LL, Harel-Bellan., A, *Histone acetyltransferases and deacetylases in the control of cell proliferation and differentiation.* Adv Cancer Res, 2002. **86**: p. 41-65.
226. Armstrong., S., Golub., TR, Korsmeyer., SJ,, *MLL-rearranged leukemias: insights from gene expression profiling.* Semin Hematol, 2003. **40**: p. 268-273.
227. Dawson., M., Bannister., AJ, Göttgens., B, et al, *JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin.* Nature., 2009. **461**: p. 819-22.
228. Lakshmikuttyamma., A, Takahashi., N, Pastural., E, *RIZ1 is potential CML tumor suppressor that is down-regulated during disease progression.* Journal of Hematology & Oncology, 2009. **2**.
229. Varambally., S., Dhanasekaran., SM, Zhou., M, et al, *The polycomb group protein EZH2 is involved in progression of prostate cancer.* Nature, 2002. **419**: p. 624-629.
230. Creusot F, AG., and Christman JK, *Inhibition of DNA methyltransferase and induction of Friend erythroleukemia cell differentiation by 5-azacytidine and 5-aza-2'-deoxycytidine.* J Biol Chem, 1982. **257**: p. 2041-2048.
231. Jones PA, and Taylor., S., *Hemimethylated duplex DNAs prepared from 5-azacytidine-treated cells* Nucleic Acids Research, 1981. **9**: p. 2933-2947.
232. Thomas., NSB, D.L., Linch DC and Lowenberg B. *Cell Cycle regulation Part 1: Normal Hematopoieses* Textbook of Malignant Hematology, 1999(Taylor and Francis): p. 33-64.
233. Ubersax JA, W.E., Quang PN, et al, *Targets of the cyclin-dependent kinase Cdk1.* Nature, 2003. **425**: p. 859-864
234. Morgan DO, *Cyclin-dependant kinases: Engines, Clocks, and Microprocessors.* Annual Review of Cell and Developmental Biology 1997. **13**: p. 261-29.
235. Mailand N, and Diffley J. *CDKs Promote DNA Replication Origin Licensing in Human Cells by Protecting Cdc6 from APC/C-Dependent Proteolysis.* Cell, 2005. **122**: p. 915–926.
236. Morgan., DO, *The cell cycle Principles of Control.* 2007(New Science Press).
237. Reynisdóttir I, P.K., A Iavarone A, et al., *Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta.* Genes Dev, 1995. **9**: p. 1831-1845.
238. Bork P, *Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally?* Proteins, 1994. **17**: p. 363-74.
239. Staller P, PK., Kiermaier A, et al, *Repression of p15INK4b expression by Myc through association with Miz-1.* Nat.Cell Biol, 2001. **3**: p. 392-399.
240. Seoane J, PC., Staller P, et al, *TGFbeta influences Myc, Miz-1 and Smad to control the CDK inhibitor p15INK4b.* Nat.Cell Biol, 2001. **3**: p. 400-408.
241. Shi Y, M.J., *Mechanisms of TGF-beta signaling from cell membrane to the nucleus.* Cell, 2003. **113**: p. 685-700.
242. Dao MO, T.N., and Nolta JA,, *Reduction in levels of the cyclin-dependent kinase inhibitor p27kip-1 coupled with transforming growth factor β neutralization induces cell-cycle entry and increases retroviral transduction of primitive human hematopoietic cells* Proc. Natl. Acad. Sci 1998. **95**: p. 13006-13011.
243. Myatt SS, a.L.E., *Promiscuous and lineage-specific roles of cell cycle regulators in haematopoiesis.* Cell Division, 2007. **2**.
244. Hirai H, R.M., Kato J,et al, *Novel INK4 Proteins, p19 and p18, Are Specific Inhibitors of the Cyclin D-Dependent Kinases CDK4 and CDK6.* Molecular and Cellular Biology, 1995. **15**: p. 2672-2681.
245. Bates S, P.D., Bonetta L, et al, *Absence of cyclin D/cdk complexes in cells lacking functional retinoblastoma protein.* Oncogene, 1994. **9**: p. 1633-1640.
246. Gartel., AL, and, Radhakrishnan., SK, *Lost in Transcription: p21 Repression, Mechanisms, and Consequences* Cancer Research 2005. **65**: p. 3980-3985.
247. Ando T, Kawabe.T., Ohara H, et al, *Involvement of the interaction between p21 and proliferating cell nuclear antigen for the maintenance of G2/M arrest after DNA damage.* J. Biol. Chem., 2001. **276**: p. 42971-42977.
248. Rodriguez R, and Meuth., M., *Chk1 and p21 Cooperate to Prevent Apoptosis during DNA Replication Fork Stress.* Mol Biol Cell, 2006. **17**: p. 402-412.
249. Zender NS , S.F., Geffers R, et al, *Argyrin A Reveals a Critical Role for the Tumor Suppressor Protein p27kip1 in Mediating Antitumor Activities in Response to Proteasome Inhibition* Cancer Cell 2008. **14**: p. 23 - 35.
250. Olashaw N, B.T.a.P.W., *Cell Cycle Control: A Complex Issue.* Cell Cycle, 2004. **3**: p. 263-4.

251. Aggerholm., A., Guldberg., P, Hokland., M, Hokland., P, , *Extensive intra- and interindividual heterogeneity of p15INK4B methylation in acute myeloid leukemia.* Cancer Res, 1999. **59**: p. 436-441.
252. Raj., K., John., A, Ho., A, et al, *CDKN2B methylation status and isolated chromosome 7 abnormalities predict responses to treatment with 5-azacytidine.* Leukemia, 2007. **21**: p. 1937-1944.
253. Jamieson AC, M.J., and Pabo CO., *Drug discovery with engineered zinc-finger proteins.* 2003. **2**.
254. Jamieson., A., Miller., JC, and Pabo., CO,, *Drug discovery with engineered zinc-finger proteins.* Nature Reviews Drug Discovery, 2003. **2**: p. 361-368.
255. Christy B, and Nathans., D., *DNA binding site of the growth factor-inducible protein Zif268.* Proc. Natl. Acad.Sci, 1989. **86**: p. 8737-8741.
256. Mandell JG, a.B.C., *Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases.* Nucleic Acids Research 2006. **34**: p. 516-523.
257. Liu Q, X.Z., Zhong X, Case CC, *Validated zinc finger protein designs for all 16 GNN DNA triplet targets.* J Biol Chem, 2002. **277**: p. 3850-6.
258. Dreier B, B.R., Segal DJ, Flippin JD, Barbas CF 3rd, *Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors.* J Biol Chem., 2001. **276**: p. 29466-78.
259. Dreier B, S.D., Barbas CF 3rd. *Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains.* J Mol Biol, 2000. **303**: p. 489-502.
260. Sera T, U.C., *Rational design of artificial zinc-finger proteins using a non-degenerate recognition code table..* Biochemistry, 2002. **41**: p. 7074-7081.
261. Dreier., B., Segal., DJ, and Barbas., CF, 3rd,, *Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains* J Mol Biol, 2000. **303**: p. 489-502.
262. Pavletich., N., and Pabo., CO,, *Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers.* Science, 1993. **261**: p. 1701-1707.
263. Clemens., K., Liao., X, Wolf., V, et al,, *Definition of the binding sites of individual zinc fingers in the transcription factor IIIA-5S RNA gene complex.* Proc Natl Acad Sci U S A. 1992. **89**: p. 10822-6.
264. Bibikova M, C.D., Segal DJ,et al, *Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases.* Mol Biol Cell, 2001. **21**: p. 289–297.
265. Liu., Q., Segal., DJ, Ghiara., JB, Barbas., CF, 3rd., *Design of polydactyl zinc-finger proteins for unique addressing within complex genomes.* Proc Natl Acad Sci U S A. , 1997. **94**: p. 5525-30.
266. Xu GL, B.T.R., *Cytosine methylation targetted to pre-determined sequences.* Nat Genet, 1997 **17**: p. 376-8.
267. Minczuk M, P.M., Kolasinska P, Murphy MP, and Klug A., *Sequence-specific modification of mitochondrial DNA using a chimeric zinc finger methylase.* PNAS, 2006. **103**: p. 19689–19694.
268. Li F, P.M., Minczuk M, et al, *Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes.* Nucleic Acids Research, 2007. **35**: p. 100–112.
269. Smith., AE, Paul., H, Bannister., AJ, Kouzarides T & Ford KG, *Heritable gene repression through the action of a directed DNA methyltransferase at a chromosomal locus.* The Journal of Biological Chemistry, 2008. **283**: p. 9878-9885.
270. (http://www.scripps.edu/mb/barbas/zfdesign/zfdesignhome.php).
271. Smith, A., Ford.,K., *Specific targeting of cytosine methylation to DNA sequences in vivo.* Nucleic Acids Research, 2007. **35**: p. 740-754.
272. Staller., P., Peukert., K, Kiermaier., A, et al, , *Repression of p15INK4b expression by Myc through association with Miz-1.* Nat Cell Biol., 2001. **3**: p. 392-9.
273. McCarthy., N., *Sssh, we are about to begin.* Nature Reviews Cancer, 2006. **418**.
274. Narita., M., Nunez,. S, Heard., E, et al, , *Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence.* Cell. , 2003. **113**: p. 703-16.
275. Peters., G., *An INKlination for epigenetic control of senescence.* Nat Struct Mol Biol., 2008. **15**: p. 1133-4.
276. Amati., B., *Integrating Myc and TGF-beta signalling in cell-cycle control.* Nat Cell Biol., 2001. **3**: p. 112-3.
277. Seoane., J., Pouponnot., C, Staller., P, et al, *TGFbeta influences Myc, Miz-1 and Smad to control the CDK inhibitor p15INK4b.* Nat Cell Biol. , 2001: p. 400-8.

278. Basu., S., Liu., Q, Qiu., Y, Dong., F, *Gfi-1 represses CDKN2B encoding p15INK4B through interaction with Miz-1.* Proc Natl Acad Sci U S A. , 2009. **106**: p. 1433-8. .
279. Heldin CH, M.K., and Dijke P. *TGF-beta signalling from cell membrane to nucleus through SMAD proteins.* Nature, 1997. **390**: p. 465-471.
280. Hu., C., Chang., TY, Cheng., CC, et al, *Snail associates with EGR-1 and SP-1 to upregulate transcriptional activation of p15INK4b.* FEBS J. 2010: p. 1202-18.
281. Schmidt., M., Bies., J, Tamura., T, et al, *The interferon regulatory factor ICSBP/IRF-8 in combination with PU.1 up-regulates expression of tumor suppressor p15(Ink4b) in murine myeloid cells.* Blood., 2004. **103**: p. 4142-9.
282. Kollhoff KA, B.A., Stubbs L., *Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3.* Hum Mol Genet., 2003. **12**: p. 233-45.
283. KH, K., *Methylation-sensitive DNA binding by v-myb and c-myb proteins.* Oncogene, 1993. **8**: p. 111-5.
284. Mandell., J., Barbas., CF 3rd., *Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases.* Nucleic Acids Res. 2006: p. (Web Server issue):W516-23.
285. Li JM, N.M., Chandrasekharan S et al. , *Transforming growth factor beta activates the promoter of cyclin-dependant kinase inhibitor p15INK4B through an Sp1 consensus site.* J Biol Chem, 1995. **270**: p. 26750-26753.
286. Stemmer., W., Crameri., A, Ha., KD, et al., *Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides.* Gene, 1995. **164**: p. 49-53.
287. Fried., M., Crothers., DM., *Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis.* Nucleic Acids Res., 1981. **9**: p. 6505-2.
288. Garner., M., Revzin., A., *A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system.* Nucleic Acids Res., 1981. **9**: p. 3047-60.
289. Wolfe, S., Nekludova, L., and Pabo., CO. *DNA Recognition by Cys2His2 Zinc Finger Proteins.* Annu. Rev. Biophys. Biomol. Struct, 1999. **3**: p. 183-212.
290. McNamara., A., Hurd., PJ, Smith., AE, Ford KG., *Characterisation of site-biased DNA methyltransferases: specificity, affinity and subsite relationships.* Nucleic Acids Res. , 2002. **30**: p. 3818-30.
291. Li., F., Papworth., M, Minczuk., M, et al., *Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes.* Nucleic Acids Research., 2007. **35**: p. 100-112
292. Prendergast., F., and Mann., K., *Chemical and physical properties of aequorin and the green fluorescent protein isolated from Aequorea forskålea* Biochemistry, 1978. **17**: p. 3448-53.
293. Sasaki., S., Kitagawa.,Y., Sekido., Y, et al,. *Molecular processes of chromosome 9p21 deletions in human cancers. .* Oncogene, 2003: p. 3792-3798.
294. Johnston., R., Pickett., SC, Barker., DL., *Autoradiography using storage phosphor technology.* Electrophoresis, 1990. **11**: p. 355-360.
295. Wolfe., S., Grant., RA, Elrod-Erickson., M, Pabo., CO,, *Beyond the "recognition code": structures of two Cys2His2 zinc finger/TATA box complexes.* Structure, 2001. **9**: p. 717-23.
296. Segal., D., Beerli., RR, Blancafort., P, et al, *Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins.* Biochemistry, 2003 **42**: p. 2137-48.
297. Isalan., M., Choo., Y and Klug., A, *Synergy between adjacent zinc fingers in sequence-specific DNA recognition.* Proc. Natl. Acad. Sci. USA, 1997. **94**: p. 5617-5621.
298. Robinson., C., and Sauer., RT., *Optimizing the stability of single-chain proteins by linker length and composition mutagenesis.* Proc Natl Acad Sci U S A., 1998. **95**: p. 5929-34.
299. Ramachandran., G., and Sasisekharan., V., *Conformation of polypeptides and proteins.* Adv Protein Chem., 1968: p. 283-438.
300. Peng., B., Fleming., JB, Breslin., T, et al., *Suppression of tumorigenesis and induction of p15(ink4b) by Smad4/DPC4 in human pancreatic cancer cells.* Clin Cancer Res., 2002. **8**: p. 3628-38.
301. Li., J., Nichols., Chandrasekharan., M, et al, *Transforming growth factor beta activates the promoter of cyclin-dependent kinase inhibitor p15INK4B through an Sp1 consensus site.* The Journal of Biological Chemistry 1995. **270**: p. 26750-26753.

302. Larsen, F., Gundersen,G., Lopez, R., and Prydz, H., *CpG islands as gene markers in the human genome.* Genomics 1992. **13**: p. 1095-1107.
303. Guenther, M., Levine, SS.,Boyer, LA., et al., *A chromatin landmark and transcription initiation at most promoters in human cells.* Cell, 2007. **130**: p. 77-88.
304. Prestridge, D., and C. Burks, C. *The density of transcriptional elements in promoter and non-promoter sequences.* Hum. Mol. Genet., 1993. **2**: p. 1449-1453.
305. Thomson, J., Skene, PJ., Selfridge,J., et al., *CpG islands influence chromatin structure via the CpG-binding protein Cfp1.* Nature, 2010. **464|**: p. 1082-1087.
306. Nguyen, C., Liang, G., Nguyen, TT., et al., *Susceptibility of nonpromoter CpG islands to de novo methylation in normal and neoplastic cells.* J. Natl. Cancer Inst, 2001. **93**: p. 1465-1472.
307. Watt., F., and Molloy., PL., *Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter.* Genes Dev. 1988. **2**: p. 1136-43.
308. Nan, X., Ng,HH.,Johnson,CA., et al., *Transcriptional repression by the methyl- CpG-binding protein MeCP2 involves a histone deacetylase complex.* Nature Cell Biology, 1998. **393**: p. 386-389.
309. Jones, P., Veenstra, GJ., Wade, PA., et al., *Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription.* Nat. Genet. , 1998. **19**: p. 187-191.
310. Greger, V., Passarge, E., Hopping, W., et al., *Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma.* Hum. Genet, 1989. **83**: p. 155 -158.
311. Herman, J., Latif, F., Weng, Y., et al., *Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma.* Proc. Natl. Acad. Sci. USA, 1994. **91**: p. 9700 - 9704.
312. Merlo, A., Herman,JG., Mao, L., et al. *5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers.* Nat Med. 1995. **1**: p. 686-92.
313. Herman, J., Merlo A., Mao, L., et al.,, *Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers.* Cancer Res., 1995. **55**: p. 4525-30.
314. Gonzalez-Zulueta, M., Bender, CM., Yang, AS., et al., *Methylation of the 5' CpG island of the p16/CDKN2 tumor suppressor gene in normal and transformed human tissues correlates with gene silencing.* Cancer Res., 1995. **55**: p. 4531-5.
315. Esteller, M., Corn, PG., Baylin, SB., Herman, JG., *A gene hypermethylation profile of human cancer.* Cancer Res., 2001. **61**: p. 3225-9.
316. Esteller, M., Fraga, MF., Guo, M., et al., *DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis.* Hum Mol Genet., 2001 **10**: p. 3001-7.
317. Robertson, K., Uzvolgyi, E., Liang, G., et al. *The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors* Nucleic Acids Res., 1999. **27**: p. 2291-2298.
318. Mizuno, S., Chijiwa, T., Okamura, T., et al, *Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia.* Blood, 2001. **97**: p. 1172-1179.
319. Swain, J., Stewart, TA., Leder, P., *Parental legacy determines melhylation and expression of an autosomal transgene: a molecular mechanism for parental imprinting.* Cell, 1987. **50**: p. 719-727.
320. Mohandas, T., Sparkes, RS., Shapiro, LJ., *Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation.* Science, 1981. **211**: p. 393-6.
321. Warnecke, P., and Clark,SJ., *DNA methylation profile of the mouse skeletal alpha-actin promoter during development and differentiation.* Molecular and Cellular Biology, 1999. **19**: p. 164-172.
322. Graff, J., Herman, JG., Myohanen, S., Baylin, SB., Vertino, PM. *Mapping Patterns of CpG Island Methylation in Normal and Neoplastic Cells Implicates Both Upstream and Downstream Regions in de novo Methylation.* J. Biol. Chem., 1997. **272**: p. 22322-22329.
323. Song, J., Stirzaker, C., Harrison, J., et al., *Hypermethylation trigger of the glutathione-S-transferase gene (GSTP1) in prostate cancer cells.* Oncogene, 2002. **21**: p. 1048-1061.

324. Macleod, D., Charlton, J., Mullins,J., et al., *Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island.* Genes Dev., 1994 : **8**: p. 2282-2292.
325. Brandeis, M., Frank,D., Keshet,I., et al.,, *Sp1 elements protect a CpG island from de novo methylation.* Nature 1994. **371**: p. 435 - 438.
326. Jost, J.-P., Frémont, M., Siegmann, M., Hofsteenge J. *The RNA moiety of chick embryo 5-methylcytosine- DNA glycosylase targets DNA demethylation.* Nucleic Acids Res., 1997. **25**: p. 4545-4550.
327. Kass, S., Pruss, D., Wolffe, AP. *How does DNA methylation repress transcription?* Trends Genet, 1997. **13**: p. 444-449.
328. Delgado, S., Gomez, M., Bird, A., and Antequera, F., *Initiation of DNA replication at CpG islands in mammalian chromosomes.* EMBO 1998. **17**: p. 2426-2435
329. Giacca, M., Zentilin, L., Norio, P., et al., *Fine mapping of a replication origin of human DNA.* Proc Natl Acad Sci USA, 1994. **91**: p. 7119-7123.
330. Taira, T., Iguchi-Ariga., SMM., and Ariga,H., *A novel DNA replication origin identified in the human heat shock protein 70 gene promoter.* Mol Cell Biol, 1994. **14**: p. 6386-6397.
331. Vassilev, L., and Johnson, EM., *An initiation zone of chromosomal DNA replication located upstream of the c-myc gene in proliferating HeLa cells.* Mol Cell Biol 1990. **10**: p. 4899-4904.
332. Gonzalez, S., Klatt, P., Delgado, S., et al., *Oncogenic activity of Cdc6 through repression of the INK4/ARF locus.* Nature. 2006 **440**: p. 702-6.
333. Antequera, F., and Bird, A.,, *CpG islands as genomic footprints of promoters that are associated with replication origins* Curr Biol. 1999. **9**: p. 661-7.
334. Esteller, M., Silva, JM., Dominguez, G., et al. *Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors.* J. Natl. Cancer Inst., 2000. **92**: p. 564-569.
335. Herman, J., Umar, A., Polyak, K., et al. *Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 6870-6875.
336. Collins, N., Wooster, R., Stratton, MR. *Absence of methylation of CpG dinucleotides within the promoter of the breast cancer susceptibility gene BRCA2 in normal tissues and in breast and ovarian cancers.* Br. J. Cancer, 1997. **76**: p. 1150-1156.
337. Hinshelwood, R., Melki, JR., Huschtscha,LI., et al., *Aberrant de novo methylation of the p16INK4A CpG island is initiated post gene silencing in association with chromatin remodeling and mimics nucleosome positioning.* Hum. Mol. Genet. , 2009. **18**: p. 3098-3109.
338. Smith.,AE, H.Paul., Bannister AJ, Kouzarides T, and Ford KG, *Heritable Gene Repression through the Action of a Directed DNA Methyltransferase at a Chromosomal Locus.* The Journal of Biological Chemistry, 2008. **283** p. 9878-9885.
339. Li F, P.M., Minczuk M, et al, *Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes.* Nucleic Acids Research, 2007. **35**: p. 100-112.
340. Smith AE, a.F.K., *Specific targeting of cytosine methylation to DNA sequences in vivo.* Nucleic Acids Research, 2007. **35**: p. 740-754.
341. McNamara, A., Hurd PJ, Smith AE, et al., *Characterisation of site biased DNA methyltransferases: specificity, affinity and subsite relationships.* Nucleic Acids Research, 2002. **30**: p. 3818-3830.
342. Xu GL, a.B.T., *Cytosine methylation targetted to pre-determined sequences.* Nature Genetics, 1997. **17**: p. 376-8.
343. Mandell, J., Barbas, CF 3rd.,, *Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases.* Nucleic Acids Res., 2006. **34**((Web Server issue):W516-23.).
344. Jamieson AC, M.J., and Pabo CO, *Drug discovery with engineered zinc finger proteins.* Nature Reviews, 2003. **2**: p. 361-368.
345. Kim YG, C.J., and Chandrasegaran S, *Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain.* Proc. Natl. Acad. Sci. USA, 1996. **93**: p. 1156-1160.
346. Kim YS, K.J., Jung DL, et al, *Artificial Zinc Finger Fusions Targeting Sp1-binding Sites and the trans-Activator-responsive Element Potently Repress Transcription and Replication of HIV-1.* The Journal of Biological Chemistry, 2005. **280**: p. 21545-21552.
347. Feng X, B.A.a.C.S., *Precise targeted integration by a chimaeric transposase zinc-finger fusion protein.* Nucl. Acids Res, 2010. **38**: p. 1204-1216.

348. Akopian, A., He, J., Boocock, M.R., Stark,W.M. *Chimeric recombinases with designed DNA sequence recognition.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**: p. 8688-8691.
349. Tan, W., Zhu, K., Segal, D.J., Barbas III, C.F., Chow, S.A.,*Fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc finger protein E2C direct integration of viral DNA into specific sites.* J. Virol. , 2004. **78**: p. 1301-1313.
350. Kim JS, K.J., Cepek KL, et al, *Design of TATA box-binding protein zinc finger fusions for targeted regulation of gene expression.* Proc Natl Acad Sci U S A 1997. **94**: p. 3616-20.
351. Thiesen, H.J., Bellefroid, E., Revelant, O., Martial, J.A.,, *Conserved KRAB protein domain identified upstream from the zinc finger region of Kox 8. .* Nucleic Acids Res., 1991. **19**: p. 3996-4005.
352. Stege, J.T., Guan, X., Ho, T., Beachy, R.N., Barbas III, C.F., *Controlling gene expression in plants using synthetic zinc finger transcription factors.* Plant J., 2002. **32**: p. 1077-1086.
353. Bartsevich, V.V., Juliano, R.L., *Regulation of the MDR1 gene by transcriptional repressors selected using peptide combinatorial libraries.* Mol. Pharmacol. , 2000. **58**: p. 1-10.
354. Beerli, R., Segal, DJ., Dreier, B., Barbas III, C.F., *Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks.* Proc. Natl. Acad. Sci. U. S. A., 1998. **95**: p. 14628-14633.
355. Yaghmai, R., Cutting, G.R.,, *Optimized regulation of gene expression using artificial transcription factors.* Molec. Ther., 2002. **5**: p. 685-694.
356. McNamara, H.P., Smith AE, and Ford KG, *Characterisation of site-biased DNA methyltransferases: specificity, affinity and subsite* Nucl. Acids Res., 2002. **30**: p. 3818-3830.
357. Segal, D.J., Goncalves, J., Eberhardy, S., et al, *Attenuation of HIV-1 replication in primary human cells with a designed zinc finger transcription factor.* J. Biol. Chem., 2004. **279**: p. 14509-14519.
358. Reynolds, L., Ullman, C., Moore, M., Isalan, M., West, M. J., Clapham, P., Klug, A., Choo, Y., *Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors. .* Proc. Natl. Acad. Sci. U. S. A., 2003 **100**: p. 1615-1620.
359. Zhang L, S.K., Liu Q, et al, *Synthetic zinc finger transcription factor action at an endogenous chromosomal site. Activation of the human erythropoietin gene.* J. Biol. Chem., 2000. **275**: p. 33850-33860.
360. Liu, P.Q., et al., *Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A.* J. Biol. Chem., 2001. **276**: p. 11323-11334.
361. Beerli, R.R., Dreier, B., Barbas III, C.F., *Positive and negative regulation of endogenous genes by designed transcription factors.* Proc. Natl. Acad. Sci. U. S. A., 2000. **97**: p. 1495-1500.
362. Ren, D., Collingwood, T.N., Rebar, E.J., et al., *PPARgamma knockdown by engineered transcription factors: exogenous PPARgamma2 but not PPARgamma1 reactivates adipogenesis.* Genes Dev, 2002. **16**: p. 27-32.
363. Tan, S., Guschin,D., Davalos, A., et al., .. , *Zinc-finger protein-targeted gene regulation: genomewide single-gene specificity.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**: p. 11997-12002.
364. Dai, Q.H., J., Klitzman, B., et al., *Engineered Zinc Finger–Activating Vascular Endothelial Growth Factor Transcription Factor Plasmid DNA Induces Therapeutic Angiogenesis in Rabbits With Hindlimb Ischemia.* Circulation, 2004. **110**: p. 2467-2475.
365. Rebar, E., Huang, Y., Hickey, R., et al., *Induction of angiogenesis in a mouse model using engineered transcription factors.* Nature Medicine 2002. **8**
366. Xu, G.L., Bestor, T.H.,, *Cytosine methylation targetted to pre-determined sequences.* Nat. Genet, 1997. **17**: p. 376-378.
367. Carvin, C.D., Parr, R.D., Kladde, M.P.,, *Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins.* Nucleic Acids Res., 2003. **31**: p. 6493-6501.
368. Teofili, L., Morosetti, R., Martini, M., et al., *Expression of cyclin-dependent kinase inhibitor p15(INK4B) during normal and leukemic myeloid differentiation.* Exp Hematol., 2000. **28**: p. 519-526.

369. Teofili, L., Martini, M., Di Mario, A.,et al.,, *Expression of p15(ink4b) gene during megakaryocytic differentiation of normal and myelodysplastic hematopoietic progenitors.* Blood 2001. **98**: p. 495-497.
370. Aggerholm, A., Guldberg, P., Hokland, M., Hokland, P., *Extensive intra- and interindividual heterogeneity of p15 INK4B methylation in acute myeloid leukemia. .* Cancer Res 1999. **59**: p. 436-444.
371. Laura, J., Christoph, P.,, *Alterations of DNA methylation in hematologic malignancies.* Cancer Lett 2002. **185**: p. 1-12.
372. Raj, K., John, A., Ho, A., et al.,, *CDKN2B methylation status and isolated chromosome 7 abnormalities predict responses to treatment with 5-azacytidine.* Leukemia, 2007. **21**: p. 1937-44.
373. Herman JG, J.J., Merlo A et al., *Hypermethylation-associated inactivation indicates a tumor suppressor role for p15INK4B.* Cancer Res 1996. **56**: p. 722-727.
374. Herman JG, C.C., Issa JP, et al., *Distinct patterns of inactivation of p15INK4B and p16INK4A characterize the major types of hematological malignancies.* Cancer Res 1997. **57**: p. 837-841.
375. Staller P, P.K., Kiermaier,et al, *Repression of p15INK4b expression by Myc through association with Miz-1.* Nature Cell Biology 2001. **3**: p. 392-399.
376. Schwartz, B.D.a., *Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process.* Nature Immunology 2003. **4**: p. 235 - 240.
377. Seoane J, P.C., Staller P, et al, *TGF influences Myc, Miz-1 and Smad to control the CDK inhibitor p15INK4b* Nature Cell Biology, 2001. **3**: p. 400-409.
378. Pagliuca A, G.P., Lania L, *Differential role for Sp1/Sp3 transcription factors in the regulation of the promoter activity of multiple cyclin-dependent kinase inhibitor genes.* Journal of Cellular Biochemistry, 2000. **76**: p. 360-367.
379. Lima PSP, M.G., Góes de Araujo A, et al, *DNA methylation analysis of the tumor suppressor gene CDKN2B in Brazilian leukemia patients.* Genetics and Molecular Biology 2008. **31**: p. 632-638.
380. Mandell, J., Barbas, CF 3rd. *Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases.* Nucleic Acids Res., 2006. **34** p. (Web Server issue):W516-23.
381. Brizzard., B., Chubet., RG, and Vizard., DL., *Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution. .* BioTechniques, 1994. **16**: p. 730-735.
382. Hopp., T., Prickett., KS, Price., VL, et al., *A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification.* Nature Biotechnology 1988. **6**: p. 1204 - 1210.
383. Hopp, T., Prickett,KS., Price,VL., et al. *A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification.* Nature Biotechnology, 1988. **6**: p. 1204 - 1210.
384. Sharp., P.a.L., WH., *The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications.* Nucleic Acids Res., 1987. **15**: p. 1281-95.
385. Sharp, P., Li, WH., *The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications.* Nucleic Acids Res., 1987. **15**: p. 1281-95.
386. Sorensen, M., Kurkland,C., and Pedersen, S.,*Codon usage determines translation rate in Escherichia coli.* J. Mol. Biol., 1989. **207** p. 365-377.
387. Lercher, M., Urrutia,AO.,Pavlıcek,A., and Hurst, LD., *A unification of mosaic structures in the human genome.* Human Molecular Genetics, 2003. **12**: p. 2411-241.
388. Varenne, S., Buc, J., Lloubes, R., Lazdunski, C.,, *Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains.* J Mol Biol, 1984. **180**: p. 549-576.
389. Dong, H., Nilsson, L., Kurland, CG.,, *Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates.* J Mol Biol, 1996. **260**: p. 649-663.
390. Kane, J.F., *Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli* Curr. Opin. Biotechnol, 1995. **6**: p. 494-500.
391. Deana, A., Ehrlich, R., Reiss, C.,*Silent mutations in the Escherichia coli ompA leader peptide region strongly affect transcription and translation in vivo.* Nucleic Acids Res., 1998. **26**: p. 4778-4782.

392. Rudolph., C., Lausier., J.; Naundorf., S., et al., *In vivo gene delivery to the lung using polyethylenimine and fractured polyamidoamine dendrimers.* Journal of Gene Medicine 2000. **2**: p. 269-78.
393. Chan, L., Nesbeth, D., MacKey, T., et al., *Conjugation of Lentivirus to Paramagnetic Particles via Nonviral Proteins Allows Efficient Concentration and Infection of Primary Acute Myeloid Leukemia Cells.* Journal of Virology, 2005. **79**: p. 13190-13194.
394. Jaenisch., R., Bird A., *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.* Nat Genet. 2003: p. 245-54.
395. Eric U. Selker, E., Freitag., M.,Kothe., GO., et al., *Induction and maintenance of nonsymmetrical DNAmethylation in Neurospora.* PNAS, 2002. **99** p. 16485-16490.
396. Lorincz., M., Schübeler.,D, Hutchinson.,SR, et al. *DNA Methylation Density Influences the Stability of an Epigenetic Imprint and Dnmt3a/b-Independent De Novo Methylation.* Mol Cell Biol., 2002. **22**: p. 7572-7580.
397. Xiong., Z., Laird.,P., et al. *COBRA: a sensitive and quantitative DNA methylation assay. .* Nucleic Acids Research, 1997. **25**: p. 2532-2534.
398. Sato., M., Sasaki., H.,Kazui., T., et al., *Probing the chromosome 9p21 region susceptible to DNA double-strand breaks in human cells in vivo by restriction enzyme transfer.* Oncogene, 2005. **24**: p. 6108-6118.
399. Sakashita., K., Koike.,K, Kinoshita.,T, et al., *Dynamic DNA methylation change in the CpG island region of p15 during human myeloid development.* The Journal of Clinical Investigation 2001 **108**.
400. Matsuno., N., Hoshino.,K, Nanri.,T, et al., *p15 mRNA expression detected by real-time quantitative reverse transcriptase-polymerase chain reaction correlates with the methylation density of the gene in adult acute leukemia.* Leukemia Research, 2005. **29**: p. 557-564.
401. Collins., S., Gallo., RC, and Gallagher., RE, *Continuous growth and differentiation of human myeloid leukaemic cells in suspension culture.* Nature 1977. **270**: p. 347-349.
402. Zheng., X., Ravatn.,R, Lin., Y, et al.,, *Gene expression of TPA induced differentiation in HL-60 cells by DNA microarray analysis.* Nucleic Acids Res. , 2002. **30**: p. 4489-4499.
403. Chim., C., Wong., AS, and Kwong., YL., *Epigenetic inactivation of INK4/CDK/RB cell cycle pathway in acute leukemias.* Ann Hematol., 2003. **82**: p. 738-42.
404. Card., C., Wilson.,GG, Weule.,K, et al. *Cloning and characterization of the Hpall methylase gene.* Nucl. Acids Res, 1990. **18**: p. 1377-1383.
405. Matsuno N, H.K., Nanri T et al, *Transcriptional repression of the p15 gene predicts the clinical outcome of acute myeloblastic leukemia with intermediate and adverse cytogenetics.* Leukemia, 2004. **18**: p. 1146-1148.
406. Suetake., I., Shinozaki., F, Miyagawa., J, et al., *DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction.* J Biol Chem. , 2004. **279**: p. 27816-23.
407. Kornberg., R., *The location of nucleosomes in chromatin: specific or statistical.* Nature, 1981. **292**: p. 579-580.
408. Chodavarapu., R., Feng., S, Bernatavichute.,YV, et al, *Relationship between nucleosome positioning and DNA methylation.* Nature, 2010. **466|**: p. 388-392.
409. Jia., D., Jurkowska., RZ, Zhang., X, et al, , *Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation.* Nature, 2007. **449**: p. 248-251.
410. Sharma., S., De Carvalho., DD, Jeong., S, et al, *Nucleosomes Containing Methylated DNA Stabilize DNA Methyltransferases 3A/3B and Ensure Faithful Epigenetic Inheritance.* PLoS Genet., 2011 **7**(e1001286).
411. Arita., K., Ariyoshi., M, Tochio., H, et al, *Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism..* Nature, 2008. **455**: p. 818-21.
412. Bronner., C., Fuhrmann., G, Chedin., FL, et al, , *UHRF1 Links the Histone Code and DNA Methylation to Ensure Faithful Epigenetic Memory Inheritance.* Genetics & Epigenetics 2009. **2**: p. 29-36.
413. Jeong., S., Liang., G, Sharma., S, et al. *Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA.* Mol Cell Biol., 2009. **29**: p. 5366-5376.
414. Chang., Y., Sun., L, Kokura., K, et al, *MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a.* Nat Commun., 2011. **2**.

415. Shamay., M., Greenway., M, Liao., G, et al, *De Novo DNA Methyltransferase DNMT3b Interacts with NEDD8-modified Proteins.* The Journal of Biological Chemistry, 2010 **285**: p. 36377-36386.
416. Jin., B., Yao., B, Li., JL, et al, *DNMT1 and DNMT3B modulate distinct polycomb-mediated histone modifications in colon cancer.* Cancer Res., 2009 **69**: p. 7412-7421.
417. Wong., D., Foster., SA, Galloway., DA, and Reid.,BJ, *Progressive Region-Specific De Novo Methylation of the p16 CpG Island in Primary Human Mammary Epithelial Cell Strains during Escape from M0 Growth Arrest.* Mol. Cell. Biol., 1999. **19**: p. 5642-5651.
418. Christiansen., D., Andersen., MK, and Pedersen-Bjergaard., J., *Methylation of p15INK4B is common, is associated with deletion of genes on chromosome arm 7q and predicts a poor prognosis in therapy-related myelodysplasia and acute myeloid leukemia.* Leukemia., 2003. **17**: p. 1813-9.
419. Xu., Y., Hu., B, Choi., AJ, et al., *Unique DNA methylome profiles in CpG island methylator phenotype colon cancers.* Genome Res., 2012. **22**: p. 283-291.
420. Mutskov., V., and Felsenfeld., G, *Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9.* EMBO J, 2004. **23**: p. 138-149.
421. Boyes., J., and Bird., A, *Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein.* The EMBO Journal, 1992. **11**: p. 327 - 333.
422. Suetake., I., Shinozaki., F, Miyagawa., J, et al, *DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction.* J Biol Chem., 2004. **279**: p. 27816-23.
423. Jia., D., Jurkowska.,RZ, Zhang., X, et al., *Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation.* Nature 2007. **449**: p. 248-251.
424. Wienholz., B., Kareta., MS, Moarefi., AH, et al. *DNMT3L modulates significant and distinct flanking sequence preference for DNA methylation by DNMT3A and DNMT3B in vivo.* PLoS Genet, 2010. **6**: p. e1001106.
425. Stewart., C., Stuhlmann., H, Jahner., D, Jaenisch., R, *De novo methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells.* Proc Natl Acad Sci USA, 1982. **79**: p. 4098-4102.
426. Li., F., Papworth., M, Minczuk.,M et al., *Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes.* Nucleic Acids Research., 2007. **35**: p. 100-112.
427. Geiman., T., Sankpal., UT, Robertson., AK, et al, *DNMT3B interacts with hSNF2H chromatin remodeling enzyme, HDACs 1 and 2, and components of the histone methylation system.* Biochem Biophys Res Commun. 2004. **318**: p. 544-55.
428. Li., Y., Zhou., PZ, Zheng., XD, et al,, *Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair.* Nucleic Acids Res, 2007. **35**: p. 390-400.
429. Gowher., H., and Jeltsch., AJ, *Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG sites.* Mol. Biol., 2001. **309**: p. 1189-1199.
430. Carvin., C., Parr., RD, and Kladde., MP, *Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins.* Nucleic Acids Res. , 2003. **31**: p. 6493-6501.
431. Hwu., W., Lee., YM., Lee., SC, and Wang., TR., *In vitro DNA methylation inhibits FMR-1 promoter.* Biochem. Biophys. Res. Commun., 1993. **193**: p. 324-329.
432. Kass., S., Goddard., JP, and Adams., RLP, *Inactive chromatin spreads from a focus of methylation.* Mol. Cell. Biol., 1993. **13**: p. 7372-7379.
433. Bryans., M., Kass., S, Seivwright., C, and Adams., RLP *Vector methylation inhibits transcription from the SV40 early promoter.* FEBS Lett, 1992. **309**: p. 97-102.
434. Choi., Y., and Chae., CB, *Demethylation of somatic and testis-specific histone H2A and H2B genes in F9 embryonal carcinoma cells.* Mol. Cell. Biol., 1993. **13**: p. 5538-5548.
435. Balaghi., M., and Wagner., C, *DNA methylation in folate deficiency: use of CpG methylase.* Biochem. Biophys. Res. Commun., 1993. **193**: p. 1184-1190.
436. Matsuo., K., Silke.,J, Gramatikoff., K, and Schaffner., W, *The CpG-specific methylase Sssl has topoisomerase activity in the presence of Mg2+.* Nucl. Acids Res. , 1994. **22**: p. 5354-5349.
437. Brooks., A., Harkins., RN, Wang., P, Qian., HS, Liu., P, Rubanyi., GM., *Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle.* J Gene Med., 2004: p. 395-404.

438. Rathert., P., Rasko., T, Roth., et al., *Reversible Inactivation of the CG Specific SssI DNA (Cytosine-C5)-Methyltransferase with a Photocleavable Protecting Group.* ChemBioChem 2007. **8**: p. 202 - 207.
439. Turker., M., *Gene silencing in mammalian cells and the spread of DNA methylation.* 2002. **21**: p. 5388-5393.
440. Boyes., J., and Bird., A, *Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein.* The EMBO Journal, 1992. **11** p. 327 - 333.
441. Nobrega., M., Ovcharenko., I, Afzal., V, and Rubin., EM, *Scanning human gene deserts for long-range enhancers.* Science, 2003. **302**: p. 413.
442. West, A., and Fraser, P., *Remote control of gene transcription.* Hum. Mol. Genet., 2005. **14**: p. R101-R111.
443. Gaszner, M., and Felsenfeld, G., *Insulators: exploiting transcriptional and epigenetic mechanisms.* Nature Rev. Genet., 2006. **7**: p. 703-713.
444. De La Rosa-Velazquez., I., Rincon-Arano., H, Benitez-Bribiesca., L, Recillas-Targa., F, *Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF.* Cancer Res., 2007. **67**: p. 2577-2585.
445. Witcher., M., and Emerson., BM *Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary.* Mol. Cell 2009. **34**: p. 271-284.
446. Kim, T., Abdullaev., ZK, Smith., AD, et al, *Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome.* Cell, 2007. **128**: p. 1231-1245.
447. Rodriguez, C., Borgel, J., Court,F., et al., *CTCF is a DNA methylation-sensitive positive regulator of the INK/ARF locus.* Biochem. Biophys. Res. Commun., 2010 **392**: p. 129-134.
448. Brandeis M, F.D., Keshet I, Siegfried Z, Mendelsohn M, et al. *Sp1elements protect a CpG island from de novo methylation.* Nature, 1994. **371**: p. 435-438.
449. Brandeis., M., Frank., D, Keshet., I, et al., *Sp1elements protect a CpG island from de novo methylation.* Nature, 1994. **371**: p. 435-438.
450. Marin., M., Karis., A, Visser., P, et al, *Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation.* Cell 1997. **89**: p. 619-628.
451. Fan., S., Fang., F, Zhang., X, and Zhang., MQ, *Putative Zinc Finger Protein Binding Sites Are Over-Represented in the Boundaries of Methylation-Resistant CpG Islands in the Human Genome.* PLoS ONE, 2007. **2**: p. e1184.
452. Dickson., J., Gowher., H, Ruslan Strogantsev., R, et al, *VEZF1 Elements Mediate Protection from DNA Methylation.* PLoS Genet, 2010. **6**: p. 1-16.
453. Gowher., H., Stuhlmann., H, and Felsenfeld.,G,, *Vezf1 regulates genomic DNA methylation through its effects on expression of DNA methyltransferase Dnmt3b.* Genes & Dev., 2008. **22**: p. 2075-2084.
454. Wu., S., Zhang., Y., *Active DNA demethylation: many roads lead to Rome.* Nat Rev Mol Cell Biol., 2010. **11**: p. 607-20.
455. Metivier., R., Gallais., R, Tiffoche., C, et al., *Cyclical DNA methylation of a transcriptionally active promoter.* Nature, 2008. **452**: p. 45-50.
456. Murayama., A., Sakura., K, Nakama., M, et al, *A specific CpG site demethylation in the human interleukin 2 gene promoter is an epigenetic memory.* EMBO 2006. **25**: p. 1081-1092.
457. Kangaspeska., S., Stride., B, Metivier., et al. *Transient cyclical methylation of promoter DNA.* Nature 2008. **452**: p. 112-115.
458. Morgan., H., Dean.,W, Coker., HA, et al., *Activation-induced Cytidine Deaminase Deaminates 5-Methylcytosine in DNA and Is Expressed in Pluripotent Tissues.* The Journal of Biological Chemistry, 2004. **279**: p. 52353-52360.
459. Bhattacharya., S., Ramchandani., S, Cervoni., N, Szyf M. *A mammalian protein with specific demethylase activity for mCpG DNA.* Nature, 1999. **397**: p. 579-83.
460. Lewis., J., Meehan., RR, Henzel., WJ, et al., *Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA.* Cell, 1992. **69**: p. 905-14.
461. Jost., J., Siegmann.,M, Sun.,L, and Leung.,R, *Mechanisms of DNA Demethylation in Chicken Embryos purification and properties of a 5-methylcytosine -DNA glycosylase P*The Journal of Biological Chemistry, 1995. **270**: p. 9734-9739.

462. Ito., S., D'Alessio., AC, Taranova., OV, et al., *Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification.* Nature, 2010. **466**: p. 1129-1133.
463. Ito., S., Shen., L, Dai., Q, et al. *Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine.* Science, 2011. **333**: p. 1300-1303.
464. Tahiliani., M., Koh.,KP, Shen.,Y, et al. *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.* Science, 2009. **324**: p. 930-935.
465. Ficz., G., Branco.,MR, Seisenberger.,S, et al., *Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation.* Nature 2011. **473**: p. 398-402.
466. Wu., H., D'Alessio.,AC, Ito.,S, et al., *Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells.* Nature, 2011. **473**: p. 389-93.
467. He., Y., Li., BZ, Li., Z, et al., *Tet-mediated formation of 5-Carboxylcytosine and its excision by TDG in mammalian DNA.* Science 2011. **333**: p. 1303-1307.
468. Ko., M., and Rao., A,, *TET2: epigenetic safeguard for HSC.* Blood 2011. **27**: p. 4501-4503.
469. Lea., N., Orr., SJ, Stoeber., K, et al, *Commitment Point during G0→G1 That Controls Entry into the Cell Cycle.* Mol. Cell. Biol., 2003. **23**: p. 2351-2361.
470. Lin., F., Ye., BG, Shen., JZ, et al, *Detection of p15 gene methylation or deletion status in different malignant cell lines by using hemi-nested methylation specific polymerase chain reaction.* Journal of experimental hematology Chinese Association of Pathophysiology, 2007. **15** p. 382-386.
471. Aitsebaomo., J., Kingsley-Kallesen., ML, Wu., Y, et al. *Vezf1/DB1 is an endothelial cell-specific transcription factor that regulates expression of the endothelin-1 promoter.* J Biol Chem., 2001. **276**: p. 39197-205.
472. Fuxe, J., Raschperger, E., Pettersson, RF., *Translation of p15.5INK4B, an N-terminally extended and fully active form of p15INK4B, is initiated from an upstream GUG codon.* Oncogene, 2000. **19**: p. 1724-8.
473. Tsubari, M., Tiihonen, E., and Laiho, M., *Cloning and characterization of p10, an alternatively spliced form of p15 cyclin-dependent kinase inhibitor.* Cancer Res 1997. **57**: p. 2966-2973.
474. Pérez de Castro., I.B., M, Jiménez., et al., *Mouse p10, an Alternative Spliced Form of p15INK4b,Transformation Inhibits Cell Cycle Progression and Malignant.* Cancer Res, 2005. **65**: p. 3249-3256.
475. Simon., M., Köster., G, Ludwig.,M et al., *Alternative splicing of the p15 cdk inhibitor in glioblastoma multiforme.* Acta Neuropathol, 2001. **102**: p. 167-174.
476. Kinney., S., Chin.,HG, Vaisvila., R, et al, *Tissue-specific Distribution and Dynamic Changes of 5-Hydroxymethylcytosine in Mammalian Genomes.* The Journal of Biological Chemistry, 2011. **286**: p. 24685-24693.
477. Brown., S., Fraga., MF, Weaver.,ICG, et al., *Variations in DNA Methylation Patterns During the Cell Cycle of HeLa Cells.* Epigenetics 2007. **2**: p. 54-6.
478. Acuto., O., and Michel., F., *CD28-mediated co-stimulation: a quantitative support for TCR signalling.* Nat Rev Immunol, 2003. **3**: p. 939-51.
479. Ficz ., G., Branco.,MR Seisenberger., S, et al, , *Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation.* Nature, 2011. **473**: p. 398-402.
480. Mellor J., *Dynamic nucleosomes and gene transcription.* Trends Genet, 2006. **22**: p. 320-329.
481. Smith., A., Chronis., C, Christodoulakis., M, et al, *Epigenetics of human T cells during the G0→G1 transition.* Genome Res., 2009. **19**: p. 1325-1337.
482. Koh., K., Yabuuchi., A,Rao., S, et al. *Tet1 and Tet2 Regulate 5-Hydroxymethylcytosine Production and Cell Lineage Specification in Mouse Embryonic Stem Cells.* Cell Stem Cell, 2011 **8**: p. 200-213.
483. Williams., K., Christensen., J, Terndrup Pedersen., M, et al., *TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity.* Nature, 2011. **473**: p. 343-348.
484. Gu ., T., Guo., F, Yang., et al, *The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes.* Nature, 2011. **477**: p. 606-10.
485. Dickson., J., Gowher., H, Strogantsev., R, Gaszner., M, Hair., A, et al., *VEZF1 Elements Mediate Protection from DNA Methylation.* PLoS Genet, 2010. **6**: p. e1000804.

486. Gowher., H., Stuhlmann., H, and Felsenfeld., G, *Vezf1 regulates genomic DNA methylation through its effects on expression of DNA methyltransferase Dnmt3b.* Genes & Dev. , 2008. **22**: p. 2075-2084.
487. Oswald., J., Engemann., S, Lane., N, et al, *Active demethylation of the paternal genome in the mouse zygote.* Curr Biol, 2000. **10**: p. 475-478.
488. Wu., S., and Zhang., Y, *Active DNA demethylation: many roads lead to Rome.* Nature reviews Molecular Cell Biology 2010 **11**: p. 607-620.
489. Langemeijer., S., Aslanyan., MG, and Janse., JH, *TET proteins in malignant hematopoiesis.* Cell Cycle 2009. **8**: p. 4044-4048.
490. Kinney., S., Chin.,HG,Vaisvila.,R, et al, *Tissue-specific Distribution and Dynamic Changes of 5-Hydroxymethylcytosine in Mammalian Genomes.* The Journal of Biological Chemistry, 2011 **286**: p. 24685-24693.
491. Bird., A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**: p. 6-21.
492. Suzuki., M., and Bird., A,, *DNA methylation landscapes: provocative insights from epigenomics.* Nature reviews Genetics 2008. **9** p. 465-476.
493. Issa JP., *CpG island methylator phenotype in cancer.* Nat Rev Cancer, 2004. **4**: p. 988-93.
494. Smith., A., Hurd.,PJ, Bannister.,AJ, Kouzarides.,T, and Ford., KG, *Heritable Gene Repression through the Action of a Directed DNA Methyltransferase at a Chromosomal Locus.* The Journal of Biological Chemistry, 2008. **283**: p. 9878-9885.
495. Mandell., J., and Barbas., CF, III, *Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases.* Nucleic Acids Res, 2006. **34**((Web Server issue)): p. W516–W523.
496. Amati., B., *Integrating Myc and TGF-beta signalling in cell-cycle control.* Nat Cell Biol. , 2001. **3**: p. E112-3.
497. Hu., C., Chang., TY, Cheng., CC, et al., *Snail associates with EGR-1 and SP-1 to upregulate transcriptional activation of p15INK4b.* FEBS Journal 2010. **277**: p. 1202-1218.
498. Minczuk., M., Papworth., MA, Kolasinska., P, Murphy., MP, Klug., A., *Sequence-specific modification of mitochondrial DNA using a chimeric zinc finger methylase.* Proc Natl Acad Sci U S A., 2006. **103**: p. 19689-94.
499. McNamara., A., Hurd., PJ, Smith., AE, and Ford., KG. *Characterisation of site-biased DNA methyltransferases: specificity, affinity and subsite relationships.* Nucleic Acids Res. , 2002. **30**: p. 3818-3830.
500. Carvin., C., Parr., RD, and Kladde., MP., *Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins.* Nucleic Acids Res, 2003. **31**: p. 6493-6501.
501. Li., F., Papworth., M, Minczuk., M, Rohde., C, Zhang., Y, Ragozin., S, Jeltsch., A., *Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes.* Nucleic Acids Res., 2007. **35**: p. 100-12.
502. Kouzine., F., Sanford.,S Elisha-Feil.,Z, and Levens.,D,, *The functional response of upstream DNA to dynamic supercoiling in vivo.* Nature Structural & Molecular Biology 2008. **15**: p. 146 - 154.
503. Bird., A., Wolffe., AP, *Methylation-induced repression--belts, braces, and chromatin.* Cell, 1999. **99**: p. 451 - 454.
504. Rountree., M., Bachman., KE, Herman., JG, and Baylin., SB., *DNA methylation, chromatin inheritance, and cancer.* Oncogene, 2001. **20**: p. 3156 - 3165.
505. Wong., D., Foster., SA, Galloway., DA, Reid., BJ., *Progressive region-specific de novo methylation of the p16 CpG island in primary human mammary epithelial cell strains during escape from M(0) growth arrest.* Mol Cell Biol., 1999 **19**: p. 5642-51.
506. Smith., A., and Ford., KG, *Specific targeting of cytosine methylation to DNA sequences in vivo.* Nucleic Acids Research, 2007. **35**: p. 740–754.
507. Gowher., H., and Jeltsch., A., *Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases.* J Biol Chem., 2002. **277**: p. 20409-14.
508. Xu., G., and Bestor., TH,, *Cytosine methylation targetted to pre-determined sequences.* Nature Genetics, 1997. **17**: p. 376 - 378.
509. Ooi., S., Qiu., C, Bernstein., E, et al,, *DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.* Nature 2007. **448**: p. 714-717.
510. Zhang., Y., Jurkowska., R, Soeroes., S, et al., *Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail.* Nucleic Acids Res., 2010. **38**: p. 4246-53.

511. Kelsey., G., *DNA methylation: a new twist in the tail.* Cell Research, 2011. **21**: p. 1155-1156.
512. Dhayalan., A., Rajavelu.,A, Rathert., P, et al, *The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation.* The Journal of Biological Chemistry, 2010. **285** p. 26114-26120.
513. Kotake., Y., Nakagawa.,T, Kitagawa. K, et al., *Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15INK4B tumor suppressor gene.* Oncogene, 2011. **30**: p. 1956-1962.
514. Vire., E., Brenner., C, Deplus., R, et al., *The Polycomb group protein EZH2 directly controls DNA methylation. .* Nature, 2006. **439**: p. 871-4.
515. Mohammad., H., Cai., Y, McGarvey., KM, et al. *Polycomb CBX7 promotes initiation of heritable repression of genes frequently silenced with cancer-specific DNA hypermethylation.* Cancer Res, 2009. **69**: p. 6322-30.
516. Kia., S., Gorski., MM, Giannakopoulos., S, Verrijzer., CP, *SWI/SNF mediates polycomb eviction and epigenetic reprogramming of the INK4b-ARF-INK4a locus. .* Mol Cell Biol., 2008. **28**: p. 3457-64.
517. Ohm., J., McGarvey., KM, Yu., X, et al, *A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. .* Nat Genet, 2007. **39**: p. 237-42.
518. Schlesinger., Y., Straussman., R, Keshet., I, et al. *Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer.* Nat Genet, 2007. **39**: p. 232-6.
519. Yu., W., Gius., D, Onyango., P, et al., *Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA.* Nature, 2008. **451**: p. 202-20.
520. Kim., T., Abdullaev., ZK, Smith., AD, et al., *Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome.* Cell 2007. **128**: p. 1231-1245.
521. De La Rosa-Velazquez., I., Rincon-Arano., H, Benitez-Bribiesca.., L, Recillas-Targa., F, *Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF.* Cancer Res, 2007. **67**: p. 2577-2585.
522. Gebhard., C., Benner., C, Ehrich., M, et al., *General Transcription Factor Binding at CpG Islands in Normal Cells Correlates with Resistance to De novo DNA Methylation in Cancer Cells.* Cancer Res, 2010 ; . **70**: p. 1398-1407.